


```

RT mRNA.":
RL Submitted (FE8-2002) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC monophosphate and cytidine monophosphate. Does not phosphorylate
CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC as a phosphate donor. Can also phosphorylate cytidine and uridine
CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
CC methylcytidine, and N(4)-antioleycytidine.
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
-----
DR EMBL; AF237290; AAK8334.1; -
DR EMBL; AF254133; AAK49122.1; -
DR EMBL; AK022317; BAB14010.1; -
DR EMBL; AF125106; AAL75943.1; -
DR InterPro; IPR006083; PRK_UK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRfams; TIGR00235; udk; 1.
KM Transferrase; Kinase; ATP-binding.
FT NP_BIND 30 37 ATP (POTENTIAL).
FT CONFLICT 8 17 DCEBPAPAD -> GARARAGAN (IN REF. 4).
FT CONFLICT 56 57 QR -> HG (IN REF. 4).
FT CONFLICT 247 247 S -> T (IN REF. 4).
SQ SEQUENCE 277 AA; 31434 MW; AFD9BD92780CD502 CRC64;

Alignment Scores:
Pred. No.: 5,53e-88 Length: 277
Score: 1450.00 Matches: 277
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.87% Gaps: 0
DB: 1

US-09-896-522-1 (1-1624) x UCK1_HUMAN (1-277)
QY 94 ATGGCTTCGGCGGAGGCGAAGATCGGAGAGCCCGCGCGGAGCGGACCGTCCGCAC 153
DB 1 MetalaetaraiaiglygluaspagyluserProalaProgluilaaspargProhis 20
QY 154 CAGGCGCCCTTCGATAGGGGTGAGCGCGGCGACTGCGACGGGAGGAGTGCAGCGTGTG 213
DB 21 GlnargProhelaueilleglyvalSercllylthraIsaSerglylserThvalCys 40
QY 214 GAGAGATCATGAGTGTCTGGAGCAGAAACGAGGTGAAACGCGGACGCGGAGGTGTC 273
DB 41 GlulysileMetGluLeuLeuGlyGlnAsnGluValGluGlnAsglnAsglyValVal 60
QY 274 ATCTGAGCCGAGCAGAGTTTACAGGTCTTGACGCGCAGAGCAGAAAGCCGAGCTTG 333
DB 61 IleleuSerGlnaPargPheTylylValLeuThraIaGluGlnlySAlaValLeu 80
QY 334 AAAGAGATCAATTTTGAACATCCAGATCCCTTTGATATGATTGATGACAGAGACT 393
DB 81 LyselyGlnTyraInPheAspHisProAspAlaPheAspAsnAspLeuMeHisArgThr 100
QY 394 CTGAAGAACATCGTGAGAGGCGAAACCGGTGAGAGGTGCGACCTTATGATTGTCACAC 453
DB 101 LeuylAsnIleValGluGlyLysThrValGluValProThryraPheValThrHis 120

```

```

QY 454 TCAAGGTTACAGAGACCAACGCGTGTCTACCCCTGCGGAGGTGTTGTTGAGGGCATC 513
DB 121 SerArgneuProIdulThrThraValIlyProAlaAspValValLeuPheGluGlyIle 140
QY 514 TTGGTGTTCACAGCCAGAGAGATCCGAGCATGTTCCACCTGCGCCCTTCGTGACACC 573
DB 141 LeuValPheylSerGlnGluIleArgAspMePheHisIleuArgneuPheValAspThr 160
QY 574 GACTCCGAGCTCAGGCTGTCTCGAAGATTTCTCCGAGAGTGGCGGAGGAGACCTG 613
DB 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgGlyArgAspLeu 180
QY 634 GACGAGATTCGAGGAGTACACACACTTCGTGAACCGGCGCTTCGAGGAGTTCGCTG 693
DB 181 GluGlnIleuThrGlnTyraInPheValIlyProAlaPheGluGluPheCysLeu 200
QY 694 CCGACAAAGAGATGATCCGATGATCATCCACGAGAGTGGACAAATATGTTGCATC 753
DB 201 ProThryllySerTyraInAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
QY 754 AACCTGATCGTGCAGACATCCAGACATTTGTAATGTGACATCTGCAATGCAACCGA 813
DB 221 AsnIleuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIlyStrHisArg 240
QY 814 GAGAGGTTCGAATGGGCGGAGCTCAAGCGGACCTTTTCGAGCGAGGGACCACTGGG 873
DB 241 GlyIlySerAsnIlyArgSerTylylAspArgInPheSerGluProIlyAspHisProGly 260
QY 874 ATCTGATCCTTCGCAACCGTGCACATTTGATGATCCAGACGAGACCAACCCGAC 924
DB 261 MetLeuThrSerIlylAspArgSerHisIleuGlnIlySerSerArgProHis 277

RESULT 2
UCK1_MOUSE
ID UCK1_MOUSE STANDARD; PRT; 277 AA.
AC P52623;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Uridine-cytidine kinase 1 (EC 2.7.1.48) (UCK 1) (Uridine
DE monophosphokinase 1) (Cytidine monophosphokinase 1).
GN UCK1 OR UMPK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshlyuk S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Guinard P.H.,
RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Guinard P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Hailton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
SEQUENCE OF 18-277 FROM N.A.

```

RC	TISSUE=Brain;
RX	MEDLINE=97108719; PubMed=8951040;
RA	Ropp P.A., Traut T.W.;
RT	"Cloning and expression of a cDNA encoding uridine kinase from mouse
RT	brain.";
RL	Arch. Biochem. Biophys. 336:105-112(1996).
CC	-1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC	monophosphate and cytidine monophosphate. Does not phosphorylate
CC	deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC	as a phosphate donor. Can also phosphorylate cytidine and uridine
CC	nucleoside analogs such as 6-thauridine, 5-fluorouridine, 4-
CC	thiouridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC	benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
CC	methylycytidine, and N(4)-anisoylcytidine (by similarity).
CC	-1- CATALYTIC ACTIVITY: ATP + uridine = ADP + CMP.
CC	-1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC	-1- PATHWAY: Pyrimidine salvage pathway.
CC	-1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC	or send an email to license@sib-sb.ch).
CC	-----
DR	EMBL; BC025146; AAH25146.1; -.
DR	EMBL; L31783; AAB50568.1; -.
DR	MGI; MGI:98904; Umpk.
DR	InterPro; IPR006083; PRK URK.
DR	InterPro; IPR000764; Uridine_kin.
DR	Pfam; PF00485; PRK; 1.
DR	PRINTS; PR00988; URIDINKINASE.
DR	TIGRFAMs; TIGR00235; udk; 1.
KW	Transferase; Kinase; ATP-binding.
FT	NP BIND 30 37
FT	SEQUENCE 277 AA; 31068 MW; 3EBB3C4187FAB64A CRC64;

Alignment Scores:		
Pred. No.:	9.08e-80	277
Score:	1325.00	Matches: 254
Percent Similarity:	95.67%	Conservative: 11
Best local Similarity:	91.70%	Mismatches: 12
Query Match:	44.66%	Indels: 0
DB:	1	Gaps: 0

QY	94	ATGCGCTTGCGGGGAGCGGAAAGACTGCGAGAGCCCCCGCGGAGCCGACCGTCCGAC	153
Db	1	Met1aSer1aIaG1yG1yG1ySer1uSer1aIaProG1uIaAspArgProG1n	20
QY	154	CAGGAGCCCTTCGATAGGGGTGAGCGCGCGCATCGCACCGGAACTGCACCTGTGCT	213
Db	21	Pro1ArgProPheLeu1IeG1yIaSer1yG1yTr1aIaSerG1yLysSerThr1aCys	40
QY	214	GAGAAAGATCATGAGATTGCTGGGACAGAGACAGGTGGAACCGGACAGGAAGGTGTC	273
Db	41	G1u1y1IleMetG1u1eG1yIeG1nMetG1u1aIaPheArg1Arg1nArg1ysLeu1aI	60
QY	274	ATCTGAGCCAGGACAGGTTCTACAGGTCCTGACGCGCAGAGCAGAGGCCAAGGCTTGG	333
Db	61	I1eUeSerG1nAspCysPheTylLysValLeuThr1aG1uG1nLys1aIbLys1aLeu	80
QY	334	AAAGGACAGTCAATTTTGGACCATTCAGATGCTTTGATTAATGATTGATGACAGCACT	393
Db	81	LysG1yG1nTr1yAsnPheAspH1sP1oAsp1aIaPheAspAsnAspLysMetC1s1yThr	100
QY	394	CTGAGAGAACATCGGAGGAGGCAAAACGCTGAGAGGCGCCGACTTGATTTGTGTGACACAC	453
Db	101	Leu1yAsn1IeValG1uG1yLysThrValG1uValPro1nTr1yAspPheValThr1s	120

Oy	454	TCMAAGTTACAGAGACCAAGGTCGTCATCCCGGACGTCGGTTCGTTTGAGGGATC	513
Db	121	SeTtgrleuprGtGtTtTtTtValtValtYtTtPtoAlaAspValValLeuphneGtGtYtLe	140
Oy	514	TTGGTGTTCACAGCCAGAGATCCGGAGATGTTCCAGCTGGCGCTTCCTGTCAGACC	573
Db	141	LeuValRheYtRtYtRtGtGtGtTtLeatGAspMetPheHtStleuYtRtLeuPheValAspThr	160
Oy	574	GACTCCGACGTCAGGCTGTCTCCAAAGTTCTCCGGACGTGCGCCGAGGGAGGGACCTG	633
Db	161	AspSerAspValRtRtLeuSerAtGAtGtValLeuAtGAspValGtAlaGtGtYtAtGAspLeu	180
Oy	634	GAGAGATTCGTGACGCGAGTACACCACTTCGTGAAGCGGCGCTTCGAGAGATTCGTCGTG	693
Db	181	GlutInIleuThtGtInIyTtRtAlaRheValIyRtProAlaRheGtGtGtIyPheCysLeu	200
Oy	694	CCGACAAAGATATGCGCATGTGATCATCCGACAGAGATGACAAATATGTTGCCATC	753
Db	201	ProThrIyAspYtYtAlaAspValIleIlePtoArgIyAlaAspAspMetValAlaIle	220
Oy	754	AACTGATTCGTGAGGCATCCGAGACATTCGTAATGTCGATCACTGGCAATAGCCACAGA	813
Db	221	AsnIleuLeuValGtInHtStIleGtInAspRtIleLeuAsnIyAspLeuCysIyAspGtHtArg	240
Oy	814	GGAAGGTCGATGAGCGCGAGATCCAAAGCGGACCTTTTCTAGCGAGGGAGGACCACTCGAG	873
Db	241	GtYtIyPtoAsnGtIyAtGAsnHtStIyAspGtThrPhePtoGtIyPtoGtIyAspHtRtPtoGtIy	260
Oy	874	ATGTGACCTCTGGCAAAAGCGTCACATTTTGGAGTCCAGACAGACACCCGAC	924
Db	261	ValIleuAlaThrGtIyAspArgSerHtStIleuGtIleuGtSerIleSerIyPtoProHtAs	277

RESULT 3	UCCK2_HUMAN	STANDARD;	PRT;	261 AA.
AC	Q9BZX2; Q96KG5;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, last sequence update)			
DT	28-FEB-2003 (Rel. 41, last annotation update)			
DE	uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (uridine monophosphokinase 2) (Cytidine monophosphokinase 2).			
GN	UCK2 OR UMPK.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
LN	[1]			
RN	SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RX	MEDLINE=21203813; PubMed=11306702;			
RA	Van Rompay A.R., Norda A., Linden K., Johansson M., Karlsson A.;			
RT	"Phosphorylation of uridine and cytidine nucleoside analogs by two human uridine-cytidine kinases."			
RL	Mol. Pharmacol. 59:1181-1186(2001).			
RN	[2]			
RP	SEQUENCE OF 15-261 FROM N.A.			
RC	TISSUE=Fibrosarcoma;			
RX	MEDLINE=21385121; PubMed=11494055;			
RA	Kozumi K., Shimamoto Y., Azuma A., Wataya Y., Matsuda A., Saeki T.,			
RT	"Cloning and expression of uridine/cytidine kinase cDNA from human fibrosarcoma cells."			
RL	Int. J. Mol. Med. 8:273-278(2001).			
CC	-1- FUNCTION: Phosphorylates uridine and cytidine to uridine monophosphate and cytidine monophosphate. Does not phosphorylate deoxynucleosides or purine ribonucleosides. Can use ATP or GTP as a phosphate donor. Can also phosphorylate cytidine and uridine nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-chlorouridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-methylcytidine, and N(4)-antisoyleytidine.			
CC	-1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.			
CC	-1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.			
CC	-1- PATHWAY: Pyrimidine salvage pathway.			

CC -1- TISSUE SPECIFICITY: Expressed in placenta.
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF236637; AAK14052.1; -
 DR EMBL; AB062451; BAB56162.1; -
 DR Genew: HGNC:12562; UMPK.
 DR InterPro; IPR006082; PRK.
 DR InterPro; IPR006083; PRK_URK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PRO0478; PHRI.BLKINASE.
 DR PRINTS; PRO0988; URIDINKINASE.
 DR TIGRFAMs; TIGR00235; udk; 1.
 DR Transferase; Kinase; ATP-binding.
 KW NP_BIND 27 34 ATP (POTENTIAL).
 FT SEQUENCE 261 AA; 29299 MW; 71791346F091EBFD CRC64;
 SQ
 Alignment Scores:
 Pred. No.: 4.46e-54 Length: 261
 Score: 934.00 Matches: 189
 Percent Similarity: 79.41% Conservative: 27
 Best Local Similarity: 69.49% Mismatches: 40
 Query Match: 31.48% Indels: 16
 Gaps: 4
 US-09-896-522-1 (1-1624) x UCK2_HUMAN (1-261)
 QY 115 GACTGCGAGAGCCCGCGCCGCGAGCCGACCGCTCGGAC--CAGCGGCGCTTCTGATA 171
 Db 4 AspergillusnitrileusGlnA:NMH:eglnA:ntrA:nsnA:ylgY:ntrA:Prophleuile 23
 QY 172 GGGGTGAGCGCGGCGGACTGCCAGCGGGAAGTGCACCGTGTGTGAAGATCAGTGG 231
 Db 24 GlycylsergylgyltrnAla:sergylYlser:serValCysAla:ylsileValGlnleu 43
 QY 232 CTGGGACAGAAAGCGAGGAGCAGACGAGGAGGTCATCCGTCGAGCGAGACAG 291
 Db 44 LeuGyl:nsmngluValAspTyr:argGln:ysglnValIleLeu:serGlnAspSer 63
 QY 292 TTTCACAGGCTCTGACGCGAGCAGAGGCGCAAGGCTTGAAGAGACAGTACATTT 351
 Db 64 PheTyr:argValleuTntr:serGlnGln:lyal:ylal:ylal:leu:ylsgln:lnPheAsnPhe 83
 QY 352 GACCATCAGATGCTTGTATGATGATTCAGACAGAGCTTGAAGAACATCTGGAG 411
 Db 84 AspHis:ProaspAla:PheAspAsnGln:leuLeu:ylserThrleu:ylsglnIleThrGln 103
 QY 412 GGCAAAACGGGAGGCGGCGGACCTATGATTTGTGCACACCAAGTTACAGAGACC 471
 Db 104 Gly:ylsThrValGlnIle:ProVal:ylAspPheValSerHis:serArg:ylsglnGlnThr 123
 QY 472 ACGGTGTCTACCTCTGGGAGCGTGTCTGTGTTGAGGAGCATCTGTGTTCTACAGCCAG 531
 Db 124 ValThrValTyrProAlaAspValValleuPheGlnGlnIleleuAlaPheTylserGln 143
 QY 532 GAGATCCGGAGCATGTTCCACTCGGCGCTCTGTGTGACACCGACCTCCAGCGCTG 591
 Db 144 GluVal:ArgAspLeuPheGlnMet:ylsLeuPheValAspThrAspAlaAspThrArgLeu 163
 QY 592 TTCGAGAGAGTTCCTCCGGAGAGT---CGCCGAGGAGGAGCCTCGAGCGAGTTCACG 648
 Db 164 SerTyr:ArgValleuAla:ArgAspIleSerGln:ArgGln:ylArgAspLeuGlnIleLeuSer 183
 QY 649 CAGTACACCACTTCTGTGAAGCCGCGCTTCGAGAGGTTCTGCGCGACAAAGAGTAT 708
 Db 649 CAGTACACCACTTCTGTGAAGCCGCGCTTCGAGAGGTTCTGCGCGACAAAGAGTAT 708

Db 184 GlnTyrIleThrPheVal:ylsProAla:PheGlnGlnPheCysLeuProThr:ylsTyr 203
 QY 709 GCCGATGTATCATCCACGAGAGTGCACATATGTTGCCATCACTGATCGTCAG 768
 Db 204 AlaAspValIleIleProArgGlyAlaAspAsnleuValAlaIleAsnleuIleValGln 223
 QY 769 CACATCCAGAGCATCTGTGATGTGACATCTGCAATGTGACACGAGAGGTCATGGG 828
 Db 224 HisIleGlnAspIleLeuAsnGlnIle-----Gly 232
 QY 829 CGAGCTACACGCGGACCTTTTCTGTGACGAGGAGCACCCTGGAGTCTGACTGTGC 888
 Db 233 ProSerTyr:argGln:ntr-----AsnGlyCysLeuAsnGlnIlyrThrProSerArg 249
 QY 889 AACGCTACATTTGGATTCACGACAGACACCCAC 924
 Db 250 LysArgGlnAla:serGln:ylserSerSerArgProHis 261
 RESULT 4
 UCK2_MOUSE
 ID UCK2_MOUSE STANDARD; PRT; 261 AA.
 AC Q99FW9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (Uridine
 DE monophosphokinase 2) (Cytidine monophosphokinase 2).
 GN UCK2 OR UMPK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21203813; PubMed=11306702;
 RA Van Rompey A.R., Norda A., Linden K., Johansson M., Karlsson A.;
 RT "Phosphorylation of uridine and cytidine nucleoside analogs by two
 RT human uridine-cytidine kinases.";
 RL Mol. Pharmacol. 59:1181-1186(2001).
 CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
 CC monophosphate and cytidine monophosphate. Does not phosphorylate
 CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
 CC as a phosphate donor. Can also phosphorylate cytidine and uridine
 CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
 CC thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
 CC benzoylcytidine, and N(4)-anisoylcytidine (By similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF236636; AAK14052.1; -
 DR MGD; MGI:1931744; Uck2.
 DR InterPro; IPR006082; PRK.
 DR InterPro; IPR006083; PRK_URK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PRO0478; PHRI.BLKINASE.
 DR PRINTS; PRO0988; URIDINKINASE.
 DR TIGRFAMs; TIGR00235; udk; 1.
 KW Transferase; Kinase; ATP-binding.
 FT NP_BIND 27 34 ATP (POTENTIAL).
 SQ SEQUENCE 261 AA; 29404 MW; 780AA3CFF5CA8153 CRC64;

Alignment Scores:

Pred. No.: 3,39e-41 Length: 260
 Score: 738.00 Matches: 143
 Percent Similarity: 82.94% Conservative: 32
 Best Local Similarity: 67.77% Mismatches: 34
 Query Match: 24.87% Indels: 2
 Gaps: 2

US-09-896-522-1 (1-1624) x UCK_DROME (1-260)

QY 160 CCCTTCCTGATGAGGCTAGCGGGCGGCTGCTCCAGCGGGAAGTGCACCGCTGTGTGAGAG 219
 Db 27 ProthleuilegylvalalaglylthralaaserglylvserrthValCyslyslvs 46
 QY 220 ATCATGAGTTCCTGCGAGACAAGCGAGTGAACAGCGGCGGAGAGTGTGATCTCTG 279
 Db 47 IllemtguglnleuglyglmlalaglmeteraphsthrnglnaaglnvalaSerlle 66
 QY 280 AGCGAGCAGAGTTCACAGAGTCTGACGCGAGAGGAGGAGGCGGCGGCTTGAAGA 339
 Db 67 SerGlnaSerPheTyArgGluLeuThrProalaglnulysalalyalaglnlygily 86
 QY 340 CAGTACATTTTGCATCCAGATCCGATGCTTGTATATGATTTGATGACACAGACTTGAAG 399
 Db 87 LeuPheanPheasphsiproAspAlaPheasnuglnleuMetTySerThrleugln 106
 QY 400 AACATCGTGAAGGCGAAACGTTGAGAGTGCAGACTTATGATTTGTGACACACTCA-- 456
 Db 107 AsnleleuylsglyhilelylvalglulileproserTyArgThAsnSerleu 126
 QY 457 AGGTTACAGAGACACCGTGTCTACCTCGGAGCGTGTGTGTGTTGAGGCACTTG 516
 Db 127 AspPhegluanValleuValilleTyProalAspValleuPhegluglylleu 146
 QY 517 GTGTTTACACCGAGAGATCCGAGATGTTTCCACCTGCGGCGGCTTGTGAGACAGGAC 576
 Db 147 ValPheTyPheProlyslleargylleuPhehshetlyslleuPheValAspThrAsp 166
 QY 577 TCCGACGTGAGTCTCTGAGAGATTTCCGCGGAGCTG--CGCGGAGGAGGAGCTG 633
 Db 167 SerasphThrArgleuAlaArgValProArgAspIleasnuglnaargglyArgAspleu 186
 QY 634 GAGCAGATTCGACGAGCTATACCACTTGTGAGGCGGCGGCTTGTGAGACTTGCCTG 693
 Db 187 AspAlaValleuThrGlnTyMetThrPheVallyspProalPhegluglnPheCysSer 206
 QY 694 CCGAGCAAGAGTATGCCAGATGATATCCACAGAGAGAGGACATATGTTGCCATC 753
 Db 207 ProThrlyslvsPheAlaAspValilleleProArgglyAlaAspAsnThrValAlaile 226
 QY 754 AACCTGATCGTGCAGACATCCAGACATCTTG 786
 Db 227 Aspleuilevalhshisileglyglulileu 237

RESULT 6
 URL1_HUMAN STANDARD; PRT; 548 AA.
 AC Q9NMZ5; O9H322;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uridine Kinase-like 1.
 GN URK1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Matanabe K., Kumeag A., Itakura S., Yamazaki M., Taahiro H., Ota T.,
 RA Suzuki Y., Odayashi M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Isegai T., Sugano S.;
 RT "NEDO human cDNA sequencing project.";

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; Pubmed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Copley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leharvalaho M.H., Leverhna M.A., Lloyd C., Lloyd D.W., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McCormachie L.J., McElay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillips B.J.C.T., Prachalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showreen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin R., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitteker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilmsham L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AK000524; BA91230.1; -;
 DR EMBL; AL118506; CAC15497.1; -;
 DR HSP; Q26998; 1BD3.
 DR Genew; HGNC:15938; URK1.
 DR InterPro: IPR006082; PRK.
 DR InterPro: IPR006083; PRK URK.
 DR InterPro: IPR00764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PR00478; PHRIBLKINASE.
 DR PRINTS; PR00988; URIDINKINASE.
 DR TIGRFAMs; TIGR00235; udk; 1.
 KW Transferase; Kinase; ATP-binding.
 FT NP_BIND 105 112 ATP (POTENTIAL).
 FT CONFLICT 219 228 MISSING (IN REF. 2).
 FT CONFLICT 290 290 N -> D (IN REF. 2).
 SQ SEQUENCE 548 AA; 61139 MW; 0CD03697E02FE7DB CRC64;

Alignment Scores:
 Pred. No.: 7.14e-29 Length: 548
 Score: 551.50 Matches: 122
 Percent Similarity: 55.37% Conservative: 48
 Best Local Similarity: 39.74% Mismatches: 68
 Query Match: 18.59% Indels: 69
 Gaps: 8

US-09-896-522-1 (1-1624) x URL1_HUMAN (1-548)
 QY 7 TCCGCTCCGACCTCGGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 66
 Db 16 SerProthrAla-----ArgAspThrProGlyArg----- 26
 QY 67 ACCGATGCGCGGAGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGC 117

```

Db 27 -----GlnAlaGluLysSerLurThraLacysGluAapArgSerAsn 40
Qy 117 -----
Db 41 AlaGluSerLeuAparGLeuProProValGlyThrGlyArgSerProArgLysArg 60
Qy 118 -----TGGCAGAGCCCGCGCGCG----- 135
Db 61 ThrThrsSerGlnYsLysSerGluProProLeuLeuArgThrsSerLysArgThrLleTyr 80
Qy 136 GAGGCCAGCCGCTCG-----CACACGCGCGCTTCTCTG 168
Db 81 ThrAlaLeuArgProProTyrArgAsnGluHisGlyThrGlnSerLysGluAlaPheAla 100
Qy 169 ATAGGGGTAGCGCGCGCGCGCGCGCGCGAGAGTGCACCGTGTGTGAAGATCAGAG 228
Db 101 IleGlyLeuGlyGlySerAlaSerGlyLysThrThrValAlaArgMetIleIleGlu 120
Qy 229 TTGCTGGAGAGAAACGAGGTGAGAACACGCGCGAGGAGGTGTCATCCTGAGCCAGGAC 288
Db 121 AlaLeu-----AspValProTyr-----ValValLeuLeuSerMetAsp 133
Qy 289 AGGTTTACAGAGCTCTGACGCGCGAGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 348
Db 134 SerPheTyrLysValLeuThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 153
Qy 349 TTGACCATCCAGATGCTCTTGTATGATGATTTGATGACAGAGCTTGAAGAACATCGTG 408
Db 154 PheAspHisProAspAlaPheAspPheAspLeuIleSerThrLeuLysLysLys 173
Qy 409 GAGGGGAAACGGGTGAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 468
Db 174 GlnGlyLysSerValLysValProLysArgPheThrThrHisSerArgLysLysAsp 193
Qy 469 ACCAGCGGTGCTACCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 528
Db 194 TrpLysThrLeuTyrGlyAlaAsnValIleIlePheGluGlyIleMetAlaPheAlaAsp 213
Qy 529 CAGAGATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 588
Db 214 LysThrLeuGlnLeuLeuLeuAspMetLysIlePheValAspThrAspSerAspIleArg 233
Qy 589 CTGCTCGAGAGTTCCTCGCGCGCGT-----CGCGAGGAGGAGCGCGCGCGAGTTCG 645
Db 234 LeuValArgArgLeuArgArgAspIleSerGlnArgLysArgAspIleGlnGlyValIle 253
Qy 646 ACCGAGTACACCACTCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 705
Db 254 LysGlnTyrAsnLysPheValLysProSerPheAspGlnTyrIleGlnProThrMetArg 273
Qy 706 TATGCCGATGTATCATCCACGAGAGGTGAGAACATATGTTGCCATCAACCTGATCGTG 765
Db 274 LeuAlaAspIleValValProArgGlySerGlyAsnThrValAlaIleAsnLeuIleVal 293
Qy 766 CAGCAGATCCAGAGATTCG 786
Db 294 GlnHisValHisSerGlnLeu 300

```

```

RESULT 7
URL_MOUSE
ID URL_MOUSE STANDARD; PRT; 548 AA.
AC 091YF3,
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Uridine kinase-like 1.
GN URKLI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Krausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
RA Saplinton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaefer T.E.,
RA Brownstein M.J., Ueda T.B., Tohyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Maira M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -I- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: BC016535; AH16535.1;
DR MGD: MGI:1915806; 111007H10R1X.
DR InterPro: IPR006082; PRK.
DR InterPro: IPR006083; PRK URK.
DR InterPro: IPR000764; Uridine_kin.
DR Pfam: PF00485; PRK.1
DR PRINTS: PR00478; PHRIDKINASE.
DR PRINTS: PR00988; URIDINKINASE.
DR TRIPRAMS: TRIGR00235; udk.1.
DR Transferrase; Kinase; ATP-binding.
FT NP BIND 105 112 ATP (POTENTIAL).
SQ SEQUENCE 548 AA; 60841 MW; 99A5CA23EB9B8525 CRC64;

```

```

Alignment Scores:
Pred. No.: 3.01e-28 Length: 548
Score: 542.00 Matches: 116
Percent Similarity: 62.5% Conservative: 46
Best Local Similarity: 44.79% Mismatches: 79
Query Match: 18.27% Indels: 19
DB: 1 Gaps: 5

```

```

US-09-896-522-1 (1-1624) x URL_MOUSE (1-548)
Qy 44 CCGGGGAGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 88
Db 49 ProProValGlyThrGlnSerLysArgProArgLysArgThrThrsSerGlnYsSerGln 68
Qy 89 CCGAGATGCTTCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 148
Db 69 ProProLeuLeuArgThrSerLysArgThrLleTyrThrAlaGlyArgProProTyr 88
Qy 149 CGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 192
Db 89 AsnGluHisGlyThrGlnSerLysGluAlaPheAlaIleGlyLeuGlyGlySerAla 108
Qy 193 AGCGGAGTGCATCCAGTGTGTGAGAGATGATGATGATGATGATGATGATGATGATGAT 252
Db 109 SerGlyLysThrThrValAlaArgMetIleIleGluAlaLeu-----AspValPro 125
Qy 253 CAGCGGCGAGGAGGTGTCATCCTGAGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 312

```

```

Db 126 Trp-----ValValleuLeuSerMetAspSerPheTyrLysValleuThrGln 141
QY 313 GAGCAGAGAGCCAGGCTTGAAGAGACAGTACATTTTGCATCCGATGCGCTTGGAT 372
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 142 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 161
QY 373 AATGATTTGATGACAGGATCTGTGAAGAACATCTGTGAGGGCAAAACGGTGGAGTGGCG 432
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 162 PheAspLeuIleIleSerThrLeuLysLeuLysGlnGlnGlnGlnGlnGlnGln 181
QY 433 ACCTATGATTTTGGACACACTCAAGTTTACCAAGAGACACAGGTGGTCTACCTCGGAC 492
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 IleTyrAspPheMetThrHisSerTyrGlnLysAspTyrLysThrLeuTyrGlnAlaasn 201
QY 493 GTGGTCTGTGTTGAGGAGATCTGTGTTCTACAGCCAGAGATCCGGAGCATGTTCCAC 552
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 202 ValIleIlePheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 221
QY 553 CTGGGCTCTTGTGGACACCGACTCCGAGCTCAGGCTGTCTGAAAGATTTCTCCGGAC 612
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 222 MetLysIlePheValAspThrAspSerAspIleArgLeuValArgArgLeuArgAsp 241
QY 613 GTG---CCCGAGGAGGAGGAGCTGGAGCAGATCTGACGACGATACACCACTTGTGAG 669
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 242 IleSerGlnArgGlnArgAspIleGlnGlnValIleLysGlnTyrAsnLysPheValLys 261
QY 670 CCGGCTTCGAGAGATCTGCTGCTCCGACAAAGATTCGCGATGTGATCATCCACGA 729
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 262 ProIlePheAspGlnTyrIleGlnProThrMetArgLeuAlaAspIleValValProArg 281
QY 730 GGAGTGACAAATATGTTTGCATCACTGATCTGTGACGACATCCAGGACATTTCTG 786
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 282 GlySerGlnAsnThrValAlaIleAspLeuIleValGlnHisValHisSerGlnLeu 300

RESULT 8
UCL_CABEL STANDARD; PRT; 229 AA.
ID_UCL_CABEL STANDARD; PRT; 229 AA.
AC Q17413;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable uridine-cytidine kinase (EC 2.7.1.48) (UCL) (Uridine
DE monophosphokinase) (Cytidine monophosphokinase).
GN B0001.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Sims M.;
RT Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL; Z69634; CA93453.1; -.
CC PIR; T18629; T18629.
CC WormPep; B0001.4; CE05142.
CC InterPro; IPR006082; PRK.
CC InterPro; IPR006083; PRK URK.
CC InterPro; IPR00764; Uridine_kin.
CC Pfam; PF00485; PRK; 1.

```

```

DR PRINTS; PR00478; PHRIBUKINASE.
DR PRINTS; PR00988; URIDINKINASE.
KW Hypothetical protein; Transferase; Kinase; ATP-binding.
FT NP_BIND 15 22 ATP (POTENTIAL)
SQ SEQUENCE 229 AA; 26195 MW; 89CB21596A820FCC CRC64;

Alignment Scores:
Pred. No.: 2,19e-26 Length: 229
Score: 512.50 Matches: 105
Percent Similarity: 69.05% Conservative: 40
Best Local Similarity: 50.00% Mismatches: 58
Query Match: 17.27% Indels: 7
DB: 1 Gaps: 4

US-09-896-522-1 (1-1624) x UCL_CABEL (1-229)
QY 160 CCCTTCTGTATAGGGGTGAGCGGCGGACCTGCGACGGGAAAGTCGACCGTGTGAGAG 219
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 8 ProLeuLeuIleGlnValAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 27
QY 220 ATCATGAGATGCTGGGAGCAGAGAGAGTGGACAG---CCGACGCGAAGGTGTGATC 276
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 28 IleIleGlnLysLeuLysAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 47
QY 277 CTGAGCCAGAGACAGGTTCTTACAAAGTCTTGCACGACAGAGAGAGCCAGGCTTGA 336
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 48 LeuSerLeuHisSerPheTyrArgGlnLeuSerAlaGlnGlnGlnGlnGlnGlnGln 67
QY 337 GACACGATCAATTTTGAACCATCCAGATGCTTGTGATTAAGATTTGATGACAGAGACTCTG 396
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 GlyLysPheAsnLysPheAspHisProAspGlnIleAsnLysPheLeuAlaGlnThrLeu 87
QY 397 AAAAACTCTGTGAGGAGCAAAACGTTGAGAGTCCGACCATATGATTTTGTGACACACTCA 456
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 88 GlnAsnMetIleAspGlnLysThrValGlnIleProLysTyrAspMetIleThr---Ser 106
QY 457 AGGTTACAGAGACACAGGCTGTCTTACCTGCGAGCGTGTCTGTTGAGGCGATCTTG 516
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 107 SerMetAsnGlnLysThrValThrValGlnProAlaLysValIleIleIleGlnGln 126
QY 517 GTGTTCTACAGCCAGAGATCCGGGACATGTTCCACTCGCTCTTGTGACACCGAC 576
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 127 LeuLeuTyrAspGlnArgValArgLysLeuLeuSerThrLysLeuIleValGlnLys 146
QY 577 TCCGAGTCGAGGCTGCTCGAAGATTT-----CTCGGAGACGTCGCGGACGAGAG 627
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 147 AlaGlnSerArgLeuArgAsnArgLeuAlaThrTyrIleArgAspTyrHisArgAlaPro 166
QY 628 GACCTGAGAGCATTTCTGACGACGATACACCACTTGTGAAAGCGGCTTTCAGAGATTG 687
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 167 -----LeuSerIleIleArgGlnTyrThrGlnPheValLysProAlaPheGlnGln 184
QY 688 TGGCTGCGGACAAAGATTCGCCGATGTGATATCCACAGAGAGTGCACATATGTT 747
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 185 CysArgProThrLysLysTyrLysValIleIleProArgGlnAlaAspAsnHisVal 204
QY 748 GCCATCAACCTGATCGGACGACATCCAG 777
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 205 AlaThrAspLeuIleAlaLysAsnLeuGln 214

RESULT 9
URK_BACHD STANDARD; PRT; 211 AA.
ID_URK_BACHD STANDARD; PRT; 211 AA.
AC Q9KDD8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine
DE monophosphokinase).
GN URK OR BH1275.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_TaxID=8665;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Maeni N.,
RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AP001511; BAB04394.1; -.
DR PIR; C83809; C83809.
DR HAMAP; MF_00551; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00478; PHRI1KINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRPFAMs; TIGR00235; udk; 1.
DR Transferase; Kinase; ATP-binding; Complete proteome.
DR NP BIND 12 19 ATP (POTENTIAL).
SQ SEQUENCE 211 AA; 24387 MW; C2AFR2CB030520B CRC64;

Alignment Scores:
Pred. No.: 8.48e-24 Length: 211
Score: 473.00 Matches: 93
Percent Similarity: 66.51% Conservative: 50
Best Local Similarity: 43.26% Mismatches: 56
Query Match: 15.94% Indels: 16
DB: Gaps: 4

US-09-896-522-1 (1-1624) x URK_BACHD (1-211)
QY 154 CACGGCCCTTCCTCGATAGGGGTGAGCGGCGACTGCCAGCGGAAGTGCACCGTGTCT 213
DB 3 LysArgProLeilellelleglyvalAlaGlyGlyThrlYserGlyLysThrThrValAla 22
QY 214 GAGAAATCATGAGAGTTGCTGTGGACAGAAACGAGTGGAGACGGCGCAGCGAAGTGTGTC 273
DB 23 LysGluIlelPhe-----TyrGlnPheGlnGluLysSerIleVal 35
QY 274 ATCTTAGAGCCAGACAGTTCTTCAAG-----GTCTTAGCGCAGACGAAAGGCC 324
DB 36 LeuIleGluIlnAspAlaTyrTyrLysAspGlnSerGlnIleuSerLeuGluGluArgLeu 55
QY 325 AAGCGCTTGAAGAGCATGATTTTTCAGATCCAGATCCAGTTCCTTATATATGATTGATG 384
DB 56 -----GlnThrAsnTyrAspHisProLeuAlaPheAspAsnAspLeuLeu 70
QY 385 CACAGAGCTGTGAAGAATCATGTGTGAGGAGAAACGGTGGAGGTGCGACCTATGATTTT 444
DB 71 IleGluHisLeuHisSerLeuAanglyGlnAlaIleLeuLysProValTyrAspTyr 90
QY 445 GTACACATCTGAAGTTTACAGAGACCAACGGTGTCTTACCTCGGACGAGTGTCTGTTT 504
DB 91 LysLeuHisThrArgSerAanglyValIleLeuValGluProLysAspValIleIleLeu 110

```

```

QY 505 GAGGACATCTGTGTTCTTACAGCCAGAGATCCGGACATGTTCCAGCGGCTCTTC 564
DB 111 GlnGlyIleLeuLeuLeuGluAspGluArgLeuAlaGluLeuMetAspIleLysLeuPhe 130
QY 565 GTGACACCCAGCTCCAGCTCAGCTGTCTCGAAGATTCTCCGGAGCGTGC---CGA 621
DB 131 ValAspThrAspAlaAspIleArgIleIleArgArgMetValArgAspIleArgGluArg 150
QY 622 GGGAGGACCTGGAGAGATTCTGACCGCAGTACACCACTTCTGTGAACCGGCTTCGAG 681
DB 151 GlyArgThrLeuGluLeuValIleGluGlnTyrThrLysValValArgProMetHisMet 170
QY 682 GAGTTCGCTGCGCCGCAAAAGATGTGCGATGTGATCATCCACAGAGAGTGAACAAT 741
DB 171 GlnPheIleGluProThrLysArgTyrAlaAspValIleIleProGluGlyGlnAsn 190
QY 742 ATGGTTCATCATCAACTGATCGTGCAGACATCATCAGACAGATTTCTG 786
DB 191 ArgValAlaIleAspLeuMetValThrLysIleArgAlaIleIle 205

RESULT 10
URK_CLOPE
ID URK_CLOPE STANDARD; PRT; 208 AA.
AC Q8XJ16;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine
DE monophosphokinase).
GN UDK OR CPE1770.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842; Oshima K., Yamashita A.,
RA Shimizu T., Ohtani K., Hirakawa H., Ogasawara N., Hayashi H.,
RA Shiba T., Ogasawara N., Hatford M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AP001511; BAB81476.1; -.
DR HAMAP; MF_00551; -.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00478; PHRI1KINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRPFAMs; TIGR00235; udk; 1.
DR Transferase; Kinase; ATP-binding; Complete proteome.
DR NP BIND 18 19 ATP (POTENTIAL).
SQ SEQUENCE 208 AA; 23884 MW; 941BC368740DDE5E CRC64;

Alignment Scores:
Pred. No.: 1.8e-23 Length: 208
Score: 468.00 Matches: 88

```

Percent Similarity: 66.51% Conservative: 55
 Best Local Similarity: 40.93% Mismatches: 56
 Query Match: 15.77% Indels: 16
 DB: 1 Gaps: 4

US-09-896-522-1 (1-1624) x URK_CLOPE (1-208)

```

QY 154 CAGGCGCCCTTCCTGATAGGGGTGAGCGCGGCACTGCACGCGGAAGTGCACCGTGTGT 213
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 2 LysArgProIlePheIleGlyIleThrGlyGlyThrGlySerGlyLysSerThrIleAla 21
QY 214 GAGAGATCATGAGTGTCTGGGACAGAAACGAGGTGAAACGCGGACCGGAGAGTGTCT 273
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 22 LysGluIleTyrArgGlnPheGlyGluAsp-----CysIleAla 34
QY 274 ATCTGACCGCAGCAGAGTCTTACAAG-----GTCTGACGCGAGACAGAGAGCC 324
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 35 MetIleGluGlnAspSerTyrTyrLysAspGlnSerHisLeuSerMetGluAspArgVal 54
QY 325 AAGGCTTGAAAGGACAGTACATTTTGCACCATCCAGATGCCCTTGTATATGATTGATG 384
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 55 LysThr-----AsnTyrAspHisProAsnIlePheAspAsnLeu 69
QY 385 CACAGACTGTGAAGACATGTGTGAGGCGAAACGGTGGAGTCCGACCTATGATTT 444
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 70 ValSerHisLeuGlnSerLeuLeuAsnGlyHisSerIleGlnLysProSerTyrAspPhe 89
QY 445 GTGACACACTCAAGTTACACAGACAGCAGGTGTCTACCTCGGAGAGTGTCTGTT 504
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 90 SerIleHisAsnArgIleGluAspThrThrLysValGlnProLysGluIleValIleVal 109
QY 505 GAGGCACTTTGTGTCTTCTACAGCCAGAGATCCGGAGACATGTTCCACCTGCGCTTTC 564
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 110 GlnGlyIleLeuIleLeuGluAspProArgIleArgIleuLeuAspIleLysIleTyr 129
QY 565 GTGACACCGACTCCGACGCTGCTGTGAGAGTGTCTCCGCGACG---CGCCGA 621
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 130 ValAspThrAspAlaAspValArgIleIleArgArgMetValArgAspIleAsnGluArg 149
QY 622 GGGAGGACCTTGAGCAGATCTTGACGACGATACACACCTTCGGAAGCGGCTTCGAG 681
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 150 GlyArgThrMetGluSerValIleAsnGlnTyrLeuAsnValValLysProMetHisAsn 169
QY 682 GAGTTCGCTGCGGACAAAGATATGCCAGTGTGATCCGACAGAGATGAGCAAT 741
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 170 GlnPheThrGlnProThrLysPheAlaAspIleIleIleProGlnGlyGlnHisAsn 189
QY 742 ATGCTGCATCAACCTGATGTCGACACATCCAGACATTTG 786
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 190 LysValAlaIleAspIleIleValAlaLysIleLysGluValLeu 204

RESULT 11
URK_LACLA STANDARD; PRT; 206 AA.
ID URK_LACLA
AC Q9CF21;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine
DE monophosphokinase).
GN URK OR L1660.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxId=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Manger S., Jallion O., Malarme K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis sep. lactis IL1403.";
RL Genome Res. 11:731-753(2001).

```

CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE006396; AAK05758.1; -.
 DR PIR; D86832; D86832.
 DR HAMAP; MF_00551; -; 1.
 DR InterPro; IPR006082; PRK.
 DR InterPro; IPR006083; PRK_URK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PR00478; PHRBLKINASE.
 DR PRINTS; PR00988; URIDINKINASE.
 DR TIGRPFAM; TIGR00235; udk; 1.
 KW Transferase; Kinase; ATP-binding; Complete proteome.
 FT NP BIND 11
 FT ATP 18
 SQ SEQUENCE 206 AA; 23740 MW; 82E38F052037DBEE CRC64;

Alignment Scores:

Pred. No.:	Length:	Matches:	Mismatches:	Indels:	Gaps:
1.5e-22	206	92	454.00	92	5
Score:		454.00			
Percent Similarity:		62.50%			
Best Local Similarity:		42.59%			
Query Match:		15.30%			

US-09-896-522-1 (1-1624) x URK_LACLA (1-206)

```

QY 154 CAGGCGCCCTTCCTGATAGGGGTGAGCGCGGCACTGCACGCGGAAGTGCACCGTGTGT 213
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 2 LysLysThrLeuIleIleGlyValThrGlyGlySerAlaSerGlyLysThrValSer 21
QY 214 GAGAGATCATGAGTGTCTGGGACAGAAACGAGGTGAAACGCGGACCGGAGAGTGTCT 273
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 22 HisAlaIleLeuGlnThrPhe---SerAsnGlu-----ArgIleAla 34
QY 274 ATCTGACCGCAGCAGTCTTACAAG-----GTCTGACGCGAGACAGAGAGCC 324
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 35 MetIleGluHisAspSerTyrTyrLysAspGlnSerHisLeuThrPheGluGluArgThr 54
QY 325 AAGGCTTGAAAGGACAGTACATTTTGCACCATCCAGATGCCCTTGTATATGATTGATG 384
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 55 LysThr-----AsnTyrAspHisProAsnIlePheAspAsnLeu 69
QY 385 CACAGACTGTGAAGACATGTGTGAGGCGAAACGGTGGAGTCCGACCTATGATTT 444
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 70 IleAlaGlnLeuLysGluLeuGlnTyrGlyArgAlaValAlaAspIleProIleTyrAspTyr 89
QY 445 GTGACACACTCAAGTTCACAGACAGCAGGTGTCTACCTCGGAGAGTGTCTGTT 504
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 90 AlaLysHisThrArgSerGlnGlnTyrTyrArgGlnLysProValAspValLeuIleVal 109
QY 505 GAGGCACTTTGTGTCTTCTACAGCCAGAGATCCGGAGACATGTTCCACCTGCGCTTTC 564
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 110 GlnGlyIleLeuValLeuGlnAspGlnArgLysGluAspLeuMetAspIleLysIlePhe 129
QY 565 GTGACACCGACTCCGACGCTGCTGTGAGAGTGTCTCCGCGACG---CGCCGA 621
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 130 ValAspThrAspAspValArgIleIleArgArgIleArgArgAspIleGlnGluArg 149
QY 622 GGGAGGACCTTGAGCAGATCTTGACGACGATACACACCTTCGGAAGCGGCTTCGAG 681
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 150 GlyArgThrLeuAspSerValIleThrGlnTyrLeuAspAlaValLysProMetTyrHis 169

```

QY 682 GAGTTCGCTGCGGACAAAGATATGCGATGATCCAGGAGAGTGACAT 741
 Db 170 GlnPheIleGutPOThrLyAspGlyAlaAspValIleLePProGluGlyAlaSerAsn 189

QY 742 ATGTTGCGATCAACCTGATCGTGCAGACATCCAGACATTCGAT 789
 Db 190 ThrValGlyAlaPheIleIleThrThrIleLeuIleSerIleuAsn 205

RESULT 12
 URK_STAMP STANDARD; PRT; 207 AA.
 ID URK_STAMP STANDARD; PRT; 207 AA.
 AC Q99TNG;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uridine kinase (EC 2.7.1.48) (uridine monophosphokinase) (Cytidine monophosphokinase).
 GN URK OR SAV1611 OR SA1439 OR MW1561.
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
 OS Staphylococcus aureus (strain N315), and
 OS Staphylococcus aureus (strain MW2).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 NX NCBI_TaxID=158878, 158879, 196620;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mu50 / ATCC 700699, and N315;
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-O., Ito T.,
 RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hoshoyama A.,
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
 RA Kenehisa M., Yamashita A., Osima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus."
 RL Lancet 357:1225-1240 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MW2;
 RX MEDLINE=22040717; PubMed=12044378;
 RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
 RA Nagai Y., Ikawa N., Asano K., Naito T., Kuroda H., Cui L.,
 RA Yamamoto K., Hiramatsu K.;
 RT "Genome and virulence determinants of high virulence community-acquired MRSA."
 RL Lancet 359:1819-1827 (2002).
 CC -|- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -|- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -|- PATHWAY: Pyrimidine salvage pathway.
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -|- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).

CC EMBL: AP003362; BAB57773.1; -
 CC EMBL: AP003134; BAB42703.1; -
 CC EMBL: AP004827; BAB95426.1; -
 CC PIR: B89943; B89943.
 CC HAMAP: MF_00551; - 1.
 CC InterPro: IPR006082; PRK.
 CC InterPro: IPR006083; PRK.
 CC InterPro: IPR00764; Uridine_kin.
 CC Pfam: PF00485; PRK_1.
 CC PRINTS: PR00478; PHRIBIKINASE.
 CC PRINTS: PR00988; URIDINKINASE.

DR TIGRPMs; TIGR00235; udk; 1.
 KM Transferase; kinase; ATP-binding; Complete proteome.
 FT NP_BIND 11 18 ATP (POTENTIAL)
 SQ SEQUENCE 207 AA; 23505 MW; FD3A6D46E531DB8D CRC64;

Alignment Scores:
 Pred. No.: 4,32e-22 Length: 207
 Score: 447.00 Matches: 88
 Percent Similarity: 65.09% Conservative: 50
 Best Local Similarity: 41.51% Mismatches: 58
 Query Match: 15.07% Indels: 16
 DB: 1 Gaps: 4

US-09-896-522-1 (1-1624) x URK_STAMP (1-207)

QY 166 CTATATGGGTGACGGCGGCGGCGGAGTCCGAGGAGTCCGATGTAAGATCATG 225
 Db 6 ILeIIeGlyIleAlaGlySerGlySerGlyThrValThrAsnGluIleMet 25

QY 226 GAGTTCGCTGCGGACAAAGCGTGCAGACGGCGGAGGAGTGCATCCGAGCCAG 285
 Db 26 LysAsnLeu-----GluGlyHisSerValAlaLeuLeuAlaGln 38

QY 286 GACAGGTTCTAC-----AAGTCTGACGCGGACAGACGAGGCGGAGGCTTGAAA 336
 Db 39 AspTyrTyrTyrIleAspGlnGlySerGlyThrPheAspGluArgLeuGluThr----- 56

QY 337 GAGCAGTACATTTGATCCATCCATGCGCTTGTGATGATGATTTGATGACAGGACTTG 396
 Db 57 -----AsnTyrAspHisProPheAlaPheAspAsnAspLeuIleGluAsnLeu 73

QY 397 AAGAAGTCGTGAGGCGAAACCGGAGGCGGCGGACGATGATTTGACACACTCA 456
 Db 74 LysAspLeuLysAsnIleLysAlaValGluValProThrTyrAspTyrAlaSerHisThr 93

QY 457 AGGTTCCAGACAGACGCGGTGCTTACCTGCGGAGCGTGGTCTGTTGAGGCGATCTTG 516
 Db 94 ArgSerAspIleThrIleAspPheLysProLysAspValIleLeuAlaGluGlyIlePhe 113

QY 517 GTGTTCTACAGCGGAGAGATCCGGGACATGTTCCATCGCGGCTTCTGTTGACCGAC 576
 Db 114 AlaLeuGlnAsnLysValLeuAsnIleArgAspMetMetAspValIleTyrValaAspThrAsp 133

QY 577 TCCGAGTACAGGCTGCTCGAAGATGTTCCGGGACGTCGCG---CGAGGAGGAGCTG 633
 Db 134 AlaAspLeuArgIleLeuArgLeuThrArgAspThrLysGluArgLysSerMet 153

QY 634 GAGCAGATTCGACGCGAGTACACACCTTCGTGAACGCGGCTTCCGAGGAGTTCGCTG 693
 Db 154 AspSerValIleAsnGlnTyrLeuSerValValArgProMetHisAspGlnPheIleGlu 173

QY 694 CCGACAAAGATGATCCGATGATCATCCACGAGAGTGCACATATGTTGCATC 753
 Db 174 ProThrLysTyrAlaAspIleIleIleProGluGlyGlySerAsnLysValaIle 193

QY 754 AACCTATGTCGACACATCCAGACATTCGAT 789
 Db 194 AspIleMetThrThrLysIleGlnSerLeuValSer 205

RESULT 13
 URK_STAMP STANDARD; PRT; 208 AA.
 ID URK_STAMP STANDARD; PRT; 208 AA.
 AC Q8P0F8;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase).
 GN URK OR SPY18_1380.
 OS Streptococcus pyogenes (serotype M18).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NX NCBI_TaxID=186103;

Alignment Scores:	
Pred. No.:	1.34e-21
Score:	439.50
Percent Similarity:	63.01%
Best Local Similarity:	39.73%
Query Match:	14.81%
DB:	1
Length:	208
Matches:	87
Conservative:	51
Mismatches:	64
Indels:	17
Gaps:	5

```

US-09-896-522-1 (1-1624) x URK_SRP8 (1-208)
QY      154  CAGCGGCCCTTCTCTGATAGGGGTGAAGCGCGGCACATGCCAGCGGGAACTGCACCGTGTGT 213
        ::::::::::::::::::::
        3  LysLeuProLeuLeuLeuLeuLeuValThrGlyGlySerGlyGlyLysThrSerValSer 22
QY      214  GAGAAGTCAATGAGTGTGCTGGGGACAGAACGAGGTGAACGCGGCAGCGGAAGTGTC 273
        ::::::::::::::::::::
        23  ArgAlaLeuLeuAspSerPhe-----ProAsnAlaArgLeuAla 35
QY      274  ATCCTAGGCAGCAGCAGGTTCCTTAAG-----GTCTGACGGCAGACAGCAAGGCC 324
        ::::::: ||| ::::::::::::::::::::
        36  MetIleGlnHisAspSerTyrtyrtyrAspArgInserHisMetSerPheGluGluArgVal 55
QY      325  AAGGCGCTTGAAGAGACAGTCAATTTTGCATCCAGATGCCCTTGTGATTAATGATTTGATG 384
        ||| ||| ||| ::::::::::::::::::::
        56  LysThr-----AsnTyrAspHisPheLeuAlaPheAspThrAspPheMet 70
QY      385  CACAGAGACTGTGAAGACATCGTGAAGGGCAAAACGTTGAAGGTGCCAGCTATGATTTT 444
        ::::: ||||| ::::::: ||| ::::::::::::::::::::
        71  IleGlnGlnLeuTySerGlnLeuLeuAlaGlyArgProValAspIleProIleTyrAspTyr 90
QY      445  GTGACACCTCAAGAGTTTACACAGAGCCACCGGTGCTACCTCGGAGCGAGTTCGTCTT 504
        ::::::::::::::::::::
        91  LysLysHisThrArgSerAsnThrThrPheArgGlnAspProGlnAspValIleVal 110

```

ID	URL STRPY	STANDARD;	PRT;	208 AA.
AC	099Z70;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase) .			
GN	UDK OR SPY1368 OR SPYM3_1042 OR SPS0818.			
OS	Streptococcus pyogenes, and			
OS	Streptococcus pyogenes (serotype M3) .			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus .			
OX	NCBI_TaxId=1314, 198466;			
NP	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SF370 / ATCC 700294 / Serotype M1;			
RX	MEDLINE=1192684; PubMed=1196296;			
RA	Peretti J.J., Mcshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,			
RA	Primeaux C., Szarek S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,			
RA	Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,			
RA	Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;			
RT	"Complete genome sequence of an M1 strain of Streptococcus pyogenes.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:4568-4563(2001).			
NN	[2]			

RC STRAIN=MGAS315 / serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Betes S.B., Sylva G.L., Barbican K.D., Lei B., Hoff U.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "genome sequence of a serotype M3 strain of group A Streptococcus:
RT phase-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / serotype M3;
RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RA Kawazaki K., Okhashi N., Kawabata S., Yasunaga T., Hattori M.,
RA Hayashi H., Hamada S.;
RT "The genome of invasive Streptococcus pyogenes; a comparative analysis
RT of S. pyogenes SSI-1, SF370 and MGAS8232.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC
CC
CC This SWISS-PROT entry is copyrighted It is produced through a collaboration


```

QY      129 CGCGCCGGA-----GGCGACGCTCCGAC-----153
DB      40 LysThrGlyLysLysIleTyrThrLysGlyArgPro-ProTyrTyrAspLysGlyLys
QY      154 -----CAGCGCCCTTCTGATAGAGGTGAGCGCGGACCTGCCAGCGGAATGCAC 206
DB      59 sSerLeuLysHisProPheValIleGlyValCysGlyGlySerHisSerGlyLysThr
QY      207 CGTGTGTGAAGAATCATGAGATTGCTGGACAGAACAGAGTGAACACGCCGCGGAA 266
DB      79 rValAlaGlyLysIleValGluArgLeuGly-----HisIleLeuLeuThrThr
QY      267 GGTGTGATCTCTGAGCGACAGATCTTCTACAAAGTCTCTGACGCGAGAGAGGACAA 326
DB      92 pValThrIleLeuSerMetAspSerPheTyrLysValLeuThrProGluGluIleLysAl 112
QY      327 GGCCTTGAAGAAGACATGATTTTGAACATCCAGATGCCCTTGATATGATTTGATGA 386
DB      112 aAlaHisGlySerArgTyrAsnPheAspGlyProAlaHisPheAspPheAspLeuLeu 132
QY      387 CAGGACTCTGAAGACATCGTGGAGGCGAAGCGTGGAGTGGCCGACCTATGATTTGT 446
DB      132 rGluValLeuLysArgLeuArgGluGlyLysSerValAspValProValTyrAspPheAs 152
QY      447 GACACACTAAGGTTACAGACACAGCGGTCTTACCCCTGCGGACGTTGTTGTTGA 506
DB      152 nThrHisSerArgAspProAsnSerLysMetMetTyrGlyAlaAspValLeuIlePheGlu 172
QY      507 GGGGATCTTGATTTCTACAGCCAGAGATCCGGAGACATGTTCCAGCTCGCCTTTGCT 566
DB      172 uGlyIleLeuAlaPheHisAspGluArgIleLysAsnLeuMetAspMetLysValPheVal 192
QY      567 GACACACGACTCCGACGCTGTCTCGAAGGTTCTCCGGAGCTGGCC--CGAGG 623
DB      192 LAspThrAspGlyLysAspLeuArgLeuAlaArgGlyIleValArgAspValThrAspArgGlu 212
QY      624 GAGGACCTGAGAGATTCTGACGACATACACACCTGCTGTAAGCGCGCTTCGAGA 683
DB      212 YArgAspIleAspGlyIleMetGluGlnTyrPheThrPheValLysProAlaPheAspLys 232
QY      684 GTTCTGCTGCGCGCAAGAAAGATAGCCGATGTGATCATCCACGAGAGTGAACAATAT 743
DB      232 sTyrIleAlaProCysMetAspSerAlaAspLeuIleValProArgGlyGlyGluAsnAs 252
QY      744 GGTGGCATCAACCTGATCTGCGACACATC 774
DB      252 pValAlaIleAspMetIleValGlnAsnVal 262

```

RESULT 2

T21110
 hypothetical protein F19B6.1b - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: T21110
 R:Thomas, K.
 submitted to the EMBL Data Library, February 1996
 A:Reference number: Z19375
 A/Accession: T21110
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-555 <WIL>
 A/Cross-references: EMBL:Z69635; NID:G1200023; PIDN:CAA93462.1; GSPDB:GN00022; CESP:F19B6
 C/Experimental source: clone F19B6
 C/Genetics:
 A/Gene: CESP:F19B6.1b
 A/Map position: 4
 A/Introns: 35/1; 83/3; 201/3; 267/3; 395/2; 511/3

Alignment Scores:

Pred. No.: 2.84e-32 Length: 555
 Score: 561.50 Matches: 125
 Percent Similarity: 63.10% Conservative: 46
 Best Local Similarity: 46.13% Mismatches: 62

```

Query Match:      18.92%      Indels:      38
DB:              2          Gaps:          6
US-09-896-522-1 (1-1624) x T21110 (1-555)

```

```

QY      12 TCCGACCTCGGCGCTGGG--CGGCGCGCCGCGGCGCGGAGAGGCGCGGCGGAG 68
DB      54 SerProArgAlaIleGlyCysArgThrArgArgThrMetSerGlyLysArgAlaGlu 73
QY      69 CCGATCGCGGAGCGGAGCGGAGCGGAGTGGCTTGGCGGAGCGGAGGCGGAGCGGAG 128
DB      74 -----HisIleLeuLeuThrThr
QY      129 CGCGCCGGA-----GGCGACGCTCCGAC-----153
DB      80 LysThrGlyLysLysIleTyrThrLysGlyArgPro-ProTyrTyrAspLysGlyLys
QY      154 -----CAGCGCCCTTCTGATAGAGGTGAGCGCGGACCTGCCAGCGGAATGCAC 206
DB      99 sSerLeuLysHisProPheValIleGlyValCysGlyGlySerHisSerGlyLysThr 119
QY      207 CGTGTGTGAAGAATCATGAGATTGCTGGACAGAACAGAGTGAACACGCCGCGGAA 266
DB      119 rValAlaGlyLysIleValGluArgLeuGly-----HisIleLeuLeuThrThr
QY      267 GGTGTGATCTCTGAGCGACAGATCTTCTACAAAGTCTCTGACGCGAGAGAGGACAA 326
DB      132 pValThrIleLeuSerMetAspSerPheTyrLysValLeuThrProGluGluIleLysAl 152
QY      327 GGCCTTGAAGAAGACATGATTTTGAACATCCAGATGCCCTTGATATGATTTGATGA 386
DB      152 nThrHisSerArgAspProAsnSerLysMetMetTyrGlyAlaAspValLeuIlePheGlu 172
QY      387 CAGGACTCTGAAGACATCGTGGAGGCGAAGCGTGGAGTGGCCGACCTATGATTTGT 446
DB      172 rGluValLeuLysArgLeuArgGluGlyLysSerValAspValProValTyrAspPheAs 192
QY      447 GACACACTAAGGTTACAGACACAGCGGTCTTACCCCTGCGGACGTTGTTGTTGA 506
DB      192 LAspThrAspGlyLysAspLeuArgLeuAlaArgGlyIleValArgAspValThrAspArgGlu 212
QY      507 GGGGATCTTGATTTCTACAGCCAGAGATCCGGAGACATGTTCCAGCTCGCCTTTGCT 566
DB      212 uGlyIleLeuAlaPheHisAspGluArgIleLysAsnLeuMetAspMetLysValPheVal 232
QY      567 GACACCGACTCCGACGCTGTCTCGAAGGTTCTCCGGAGCTGGCC--CGAGG 623
DB      232 LAspThrAspGlyLysAspLeuArgLeuAlaArgGlyIleValArgAspValThrAspArgGlu 252
QY      624 GAGGACCTGAGAGATTCTGACGACATACACACCTGCTGTAAGCGCGCTTCGAGA 683
DB      252 YArgAspIleAspGlyIleMetGluGlnTyrPheThrPheValLysProAlaPheAspLys 272
QY      684 GTTCTGCTGCGCGCAAGAAAGATAGCCGATGTGATCATCCACGAGAGTGAACAATAT 743
DB      272 sTyrIleAlaProCysMetAspSerAlaAspLeuIleValProArgGlyGlyGluAsnAs 292
QY      744 GGTGGCATCAACCTGATCTGCGACACATC 774
DB      292 pValAlaIleAspMetIleValGlnAsnVal 302

```

RESULT 3

T18629
 hypothetical protein B0001.4 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C/Accession: T18629
 R:Sim, M.
 submitted to the EMBL Data Library, February 1996
 A:Reference number: Z18999
 A/Accession: T18629
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA

A/Residues: 1-229 <WIL>
A/Cross-references: EMBL:Z69634; PIDN:CAA93453.1; GSPDB:GN00022; CESP:B0001.4
A/Experimental source: clone B0001
C/Genetics:
A/Gene: CESP:B0001.4
A/Map position: 4
A/Introns: 21/3; 41/3; 108/2
C/Superfamily: uridine kinase

Alignment Scores:

Pred. No.:	8.04e-29	Length:	229
Score:	512.50	Matches:	105
Percent Similarity:	69.05%	Conservative:	40
Best Local Similarity:	50.00%	Mismatches:	58
Query Match:	17.27%	Indels:	7
DB:	2	Gaps:	4

US-09-896-522-1 (1-1624) x T18629 (1-229)

```
QY 160 CCCTTCCTGATAGGGGTGAGCGGCGGACATGCCAGCGGGAATGCACTGCTGTGAGAG 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8 Prolenleuileglyalaglygltthsercysgilysserthrllevalgluarg 27
QY 220 ATCATGAGTGTCTGGAGACAGACGAGTGGACAG---CGCGACGCGAAGGTGTCTATC 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28 lleilegluasnleuasnalaasnalaalysglnserglyargglnlleaprllevalhis 47
QY 277 CTGAGCCAGACAGGTTCTACAAAGTCTTGACGCGACAGACAGAAAGCCCAAGCCTTGAA 336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 48 leuserleuhsierpherfyrarggluleuseralaglululyslleualaaarglu 67
QY 337 GGCACATTAACATTTTACCATTCAGATGCTTGATTAATGATTTGATGCACAGACTCTG 396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68 glylypphenasnheaphisproaspagnlleasnphenaleuualagluthrleu 87
QY 397 AAGAAATCATGTGAGGCAAAACGCTGAGAGTGCACCACTATGATTTTGTGACACACTCA 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 88 glnasmetleaspelgylsrthvalgluileprolysrthsrpmetlethr---ser 106
QY 457 AGGTTCACAGACACACGCTGTCTACCTCGGACGCTGCTGTGTTGAGGCGATCTTG 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 107 sermetasnlglythrvalthrvalgluproalalyvallellellelelglylleu 126
QY 517 GGTGTTTACACGACGAGATCCGGGACATGTTCCACTGCGCTCTTCGTCGACACCGAC 576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 127 leuileuylrarglulargvalarglylsleuuserthrlyslenuhevalglulysasn 146
QY 577 TCCGACGTACAGCTGTCTCGAAGAGTT-----CTCCGGACGTCGCGACGAGGAG 627
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 147 alagluserrargleuargasnargleuallathrtyrillearghsprlyrhisargalapro 166
QY 628 GACCTGAGACAGATTCTGACGCGATACACCACTTCGTGAAGCCGCTTCGAGAGTTTC 687
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 -----leuserlleleargglnfyrthrgluPhvallysrproalaphnegluluphe 184
QY 688 TGGCTCCCAACAAGAAGATATGCCGATGTGATATCCCAAGAGAGTGCACATATAGTT 747
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 185 cysargprthrlysrlysrlyalaasprvallelleleproargglylaaasprasnhsval 204
QY 748 GCCATCAACCTGCTGTCGACATCTCAG 777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 205 Alathraspleuilealalyasenuleugn 214
```

RESULT 4

C83809 uridine kinase udk [imported] - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C/Accession: C83809
R/Takemi, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; WUID:20512582; PMID:11058132

A/Accession: C83809
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-211 <STO>
A/Cross-references: GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BAB04994.1; GSPDB:GN001
A/Experimental source: strain C-125
C/Genetics:
A/Gene: udk
C/Superfamily: uridine kinase

Alignment Scores:

Pred. No.:	5.34e-26	Length:	211
Score:	473.00	Matches:	93
Percent Similarity:	66.51%	Conservative:	50
Best Local Similarity:	43.26%	Mismatches:	56
Query Match:	15.94%	Indels:	16
DB:	2	Gaps:	4

US-09-896-522-1 (1-1624) x C83809 (1-211)

```
QY 154 CAGCGGCTCTCTGATAGGGGTGAGCGGCGGACATGCCAGCGGGAATGCACTGCTGT 213
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 lysargprollellellelelyalalaglygltthserglysrththrvalala 22
QY 214 GGAAGATCATGAGTGTCTGGACAGAACGAGTGGAAACAGCGGACGCGAAGGTGTC 273
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 23 lysglnlephe-----tyrglnphenaglnlysrlelleval 35
QY 274 ATCTGAGCCAGACAGGTTCTACAG-----GTCTGACGCGACAGACAGAGGCC 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36 leuileglulnaspalalyrtyrlysraspglinserglnleuserleuuglnuargleu 55
QY 325 AAGGCTTGAAGAAGACAGTACATTTTGCATCCAGATGCTTGATTAATGATTTGATG 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 56 -----glnthrasnlysrphsproleuallaPhasprasnpleuenu 70
QY 385 CACAGACCTGTAGACATCGTGAGGCAAAACGCTGAGAGTGCACGCTATGATTTT 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 71 llegluhsleuhsierleuenuenglylnalalleglulysprovaltyrspytr 90
QY 445 GTGACACTCAAGTTACACAGACACGCTGTCTACCTCGGACGCTGTTCTGTT 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 91 lysleuhslethrlysrasnlglnvalleleuvalgluprolyasprvallelleleu 110
QY 505 GAGGCGATTTGTTGTTCTACACGACGAGATCCGGACATGTTCCACTGCGCTCTTC 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111 gluglylleleuenuleuglnaspglulargleuarggluleumecAsprlleysleuphe 130
QY 565 GNGACACCGACTCGACGCTGAGGCTGTCGAGAGTTCTCGGAGCGTGGCC--CGA 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 131 valasprthraspalasprlleargllelleargargmetvalargasprllearglularg 150
QY 622 GGGAGGACCTTGAGACAGATTCTGACGCGATACACCACTTCGTGAAGCCGCTTCGAG 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151 glyargthleugluservallelelglnfyrthrllysrvalvalargpromethismet 170
QY 682 GAGTTCTGCTCCCAACAAGAAGATGTCGATGATATCCCAAGAGAGTGCACAT 741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 glnpheillegluprothrlysrargtyrllaasprvallelleleprogluglylglnasn 190
QY 742 ATGGTTGCCATCAACCTGATGTCGACGACATCCAGGACATCTG 786
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 argvalalalleasprleuvalthrlysrleleargalallelle 205
```

RESULT 5

T41020 probable uridine kinase - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C/Accession: T41020
R/Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, June 1998
A/Reference number: Z21965

A/Accession: T41020
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-454 <NR>
A/Cross-references: EMBL:AL023860; PIDN:CA19591.1; GSPDB:GN00068; SPDB:SPC162.11C
A/Experimental source: strain 972h-; cosmid c162
C/Genetics:
A/Status: SPDB:SPC162.11C
A/Map position: 3

Alignment Scores:
Pred. No.: 5,996-25 Length: 454
Score: 459.00 Matches: 92
Percent Similarity: 62.33% Conservative: 47
Best Local Similarity: 41.26% Mismatches: 76
Query Match: 15.47% Indels: 8
DB: 2 Gaps: 2

US-09-896-522-1 (1-1624) x T41020 (1-454)

QY 124 AGCCCCCGCGGAGCCGACCGCTCCGACCGGCCCTTCGTATGAGGGTGAGCGGC 183
Db 9 SerSerAnPrOthrTyGluProPrOtrPrArgValArgPheIleGlyIleAlaGly 28
184 GGCACTCCAGCGGAGAGTGACCGTGTGTGAAGAATCATGTGAGTTCGTGGACAGAAC 243
Db 29 ProSerIleSerGlyIleThrSerValAlaGluIleValIleValAlaLeuAnLeuPro 48
QY 244 GAGGTGAGACAGCGGACCGGAGAGTGATCTCTGAGCCAGAGTTCTACAGATC 303
Db 49 His-----ValValIleLeuSerLeuAnSerPheTyLeuSer 61
QY 304 CTGACGGCAGAGCAGAGGCCAAGCCCTTGAAAGACATTCATTTTTCAGATCCAGAT 363
Db 62 LeuAnAlaGluGlnIleValArgAlaPheAnAnSerPhePheAnSerProGlu 81
QY 364 GCGTTGATATGATTTGATTCACAGAGCTCTGAAGAATCCGTGAGGCGAAGACGCTG 423
Db 82 AlaIleAspTrpAspLeuLeuPheValIleLeuLeuGluLeuValGlnGlyArgVal 101
QY 424 GAGGTGCGGACCTTGTGATTTGTGACACACTCAAGGTTACAGAGACCGGTGTAC 483
Db 102 AspIleProIleTySerPheAnGlnHisAnArgLeuProGluThrAnThrLeuPhe 121
QY 484 CTTGGGACGCGTCTCTGTTGAGGCACTTGTGTTCACAGCCAGAGATCCGGAC 543
Db 122 GlyAlaSerIleIleIleLeuGlnGlyIlePheAlaLeuTyArgGluValIleArgSer 141
QY 544 ATGTTCCACCTGCGCTTCGTGACACAGCACTCCGACGCTGTCTCGAAGAGTT 603
Db 142 LeuLeuAnArgValSerValPheLeuAnSerThrAspSerAspValCysLeuSerArgArgLeu 161
QY 604 CTCGGGACGTCGC---CGAGGAGGAGCCTGAGAGATTCGACGCACTACACACC 660
Db 162 AsnArgAspIleAnTyArgGlyArgAspIleValGlyValLeuGlnGlnTyArgAnlys 181
QY 661 TTTCGTGAGCGCGCTTCGAGAGATTCCTGCTGCCGCAAGAGATTCGCGAGTATC 720
Db 182 PheValIleProSerTyArgIleAnPheValArgArgIleuSerTyThrAspLeuIle 201
QY 721 ATCCACGAGAGGAGCAATATGTTGTCATCACTGATTCGTGACAGCATTCAGAG 780
Db 202 ValProArgGlyArgAspAnLeuValAlaIleAspMetValIleAnPheIleArgArg 221
QY 781 ATTCTGAAT 789
Db 222 ThrLeuSer 224

RESULT 6
D86832
uridine kinase (EC 2.7.1.48) [imported] - Lactococcus lactis subsp. lactis (strain IL140
C/Species: Lactococcus lactis subsp. lactis
C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C/Accession: D86832
R/Bolotin, A.; Wincker, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssi
A/Reference number: A86825; WUID:21235186; PMID:1137471
A/Accession: D86832
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-206 <STO>
A/Cross-references: GB:AE005176; PID:g12724672; PIDN:AAK05758.1; GSPDB:GN00146
A/Experimental source: strain IL1403
C/Genetics:
A/Status: udk
C/Superfamily: uridine kinase
C/Keywords: phosphotransferase

Alignment Scores:
Pred. No.: 1,226-24 Length: 206
Score: 454.00 Matches: 92
Percent Similarity: 62.50% Conservative: 43
Best Local Similarity: 42.59% Mismatches: 65
Query Match: 15.30% Indels: 16
DB: 2 Gaps: 5

US-09-896-522-1 (1-1624) x D86832 (1-206)

QY 154 GAGCGCCCTTCGTATGAGGGTGAGCGGCGGACCTCCAGCGGAGATCGACCGTGT 213
Db 2 LysIleThrLeuIleIleGlyValThrGlyIleSerIleSerValSer 21
QY 214 GAGAAATGATGAGAGTTCGTGGGACAGAGAGTGAACAGCGGACGAGAGTGATC 273
Db 22 HisAlaIleLeuGlnThrhe---SerAnGln-----ArgIleAla 34
QY 274 ATCTGAGCCGAGACAGGTTCTACAG-----GTCTGACGCGACAGAGAGCC 324
Db 35 MetIleGlnHisAspertyrTyIleValSerGlnSerHisLeuThrPheGluGluArgThr 54
QY 325 AAGGCTTGAAAGACAGATCAATTTTGACATCCGATCCGATGCTTGATATGATTGATG 384
Db 55 LysThr-----AnTyArgPheHisProLeuAlaPheAspThrAspTyLeu 69
QY 385 CACAGAGCTCTGAAGAATCTGTGAGGCGCAAAACCGTGAGGTGCGCATGATTTT 444
Db 70 IleAlaGlnLeuIleValLeuGlnTyArgAlaValAlaAspIleProIleTyArgTy 89
QY 445 GTGACACACTCAAGTTACCAAGACACAGCGGTGTACTCTCGGACGTGTTCTGTT 504
Db 90 AlaIleHisThrArgSerGlnGlnTyArgGlnIleProValAspValIleVal 109
QY 505 GAGGCACTTGTGTTCTTACAGCCAGAGATCCGGAGATGTTCCACTGCGCCTTC 564
Db 110 GlnGlyIleLeuValIleuGlnAspGlnArgLeuArgAspLeuSerAspIleValIlePhe 129
QY 565 GTGACACAGCACTCCGACCTCAGAGTGTCTGGAAGAGTTCTCGGAGACGTG---CGCCGA 621
Db 130 ValAspThrAspAspValArgIleIleValGlnGlyIleArgArgAspIleGluGlnArg 149
QY 622 GGGAGGACCTGAGCAGATTCGACGCACTACACACACTTGTGAGCCGCGCTTCGAG 681
Db 150 GlyArgThrLeuAnSerValIleThrGlnTyLeuAnAspAlaValIleAspMetTyHis 169
QY 682 GAGTTCTGCTGCGGCAAGAGATTCGCGATGATATCCGACGAGAGTGCAGAT 741
Db 170 GlnPheIleGluProThrIleArgTyArgValAlaAspValIleIleProGluGlyValSerAn 189
QY 742 ATGTTGCATCAACTGATTCGTGACAGCATTCGAGACATTCGAT 789
Db 190 ThrValGlyValAspIleIleThrThrIleValSerIleLeuAn 205

RESULT 7
B89943
uridine kinase [imported] - Staphylococcus aureus (strain N315)

[illegible]

RESULT 8

B95140 uridine kinase [imported] - Streptococcus pneumoniae (strain TIGR4)

C:Species: Streptococcus pneumoniae

C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001

C:Accession: B95140

R:Reteljin, H.; Neilson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidson, J.D.; Unayam, L.A.; White, O.; Salsbery, S.L.; Lewis, M.R.; Radune, D.; Holtzaple, I.E.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: B95140

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-212 <KOR>

A:Cross-references: GB:AE005672; PIDN:AAK75315.1; PID:g14972689; GSPDB:GN00164; TIGR:SP46

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SPI208

C:Superfamily: uridine kinase

Alignment Scores:

Pred. No.:	1-71e-23	Length:	212
Score:	438.00	Matches:	88
Percent Similarity:	60.75%	Conservative:	42
Best Local Similarity:	41.12%	Mismatches:	68
Query Match:	14.76%	Indels:	16
DB:	2	Gaps:	4

US-09-896-522-1 (1-1624) x B95140 (1-212)

Oy 157 CGGCCCTTCGTATGAGGGGTGAGCGGCGGCGGACTGCCAGCGGGAAGTGCACGTGTGTAG 216

Db 4 ArgProIleIleIleGlyValThrGlyGlySerGlyGlyValThrSerValSerArg 23

Oy 217 AAGATCATGAGACTCTCTGGAGACAAACGAGTGGTGAACACCGGAGCGGGAAGGTGTCATC 276

Db 24 AAlaIleuSer-----HisPheProAspIuIySileSerMet 36

Oy 277 CTGAGCCGACGACAGGTTCTACAG-----GTCCTACGCGACGACGAGGACCAAG 327

Db 37 IleguHshApSerTyTrIyIsaSpGlnSerHsIleThrPheIuGluValGlyLys 56

Oy 328 GCCTTGAAGAAGACAGTACAAATTTTGACCATCCAGATGCTTGATATGATTTGATGCAC 387

Db 57 Thr-----AsnTyraSpHisProPheAlaPheAspThrAspLeuMetIle 71

Oy 388 AGGACTCTGAAGACATGCTGTGAGGGCAAAACGGTGGAGAGTCCGACCTATGATTTTGTG 447

Db 72 GluGlnIleIyValGluLeuIleuAlaGlyArgProValAspIleProThrIyAspTyTrThr 91

Oy 448 ACACACTCAAGATTACAGACACGACGAGCGGTCTCAACCTCCGCGACGCTGTCTTTGAG 507

Db 92 GluHshThrIyAspSerIyThrIyArgGlnGluProGlnAspValPheIleValGlu 111

Oy 508 GGCACTTGTGTGTTCTTACAGCCAGAGATCCGGAGCATGTTCCACTCGCGCTTCTCG 567

Db 112 GlyIleIeuValIleuGluAspIyAspIyArgLeuMetAspIleIySilePheVal 131

Oy 568 GACACCGCATCCGACAGCTGTCTGTGAAAGATTCTCCGGAAGCTG---CCCGAGGG 624

Db 132 AspThrAspAspAspValArgIleIleArgGlyIleIyAspMetGluIuArgGly 151

Oy 625 AGGCACTCGAGACAGATTCTGACGACATACACACCTTGTTGAAGCGGCTTCGAGAG 684

Db 152 ArgSerIeuAspSerValIleAsnGlnTyIleuGlyValValIyAspProMetTyIyHsGln 171

Oy 685 TTTCGCTCCGACAAAGAGATATGCCGATGTGATCATATCCACGAGAGAGTGCACATATG 744

Db 172 PheIleGluSerThrIyAspGlyTrIyAlaSpIleValIleProGlnGluIyAlaSerAsnThr 191

Qy 745 GTTGCCATCAACCTGATCGTGACACATCCAGACATCTTG 786
 Db 192 ValAlaIleAspLeuThrThrLysIleAlaLysIleLeu 205

RESULT 9

A99008

uridine kinase (EC 2.7.1.48) [imported] - Streptococcus pneumoniae (strain R6)

C/Species: Streptococcus pneumoniae

C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001

C/Accession: A99008

R/Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A/Reference number: A97872; MUID:21429245; PMID:11544234

A/Accession: A99008

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-212 <KOR>

A/Cross-references: GB:AE007317; PIDN:AAK9893.1; PID:G15458714; GSPDB:GN00174

C/Genetics:

A:Gene: udk

C/Superfamily: uridine kinase

C/Keywords: phosphotransferase

Alignment Scores:

Pred. No.:	1-71e-23	Length:	212
Score:	438.00	Matches:	88
Percent Similarity:	60.75%	Conservative:	42
Best Local Similarity:	41.12%	Mismatches:	68
Query Match:	14.76%	Indels:	16
DB:	2	Gaps:	4

US-09-896-522-1 (1-1624) x A99008 (1-212)

Qy 157 CGGCCCTTCTGATGAGGGGTGAGCGCGGCGACCTCCAGCGGAAGTCGACCGTGTGAG 216

Db 4 ArgProIleIleIleGlyValThrGlySerGlyGlyLysThrSerValSerArg 23

Qy 217 AAGTGTGAGTGTCTCGGACAGACAGAGTGTGAACGCGGACCGGAAGTGTATC 276

Db 24 AAlaIleUser-----HisPheProAspGluLysIleSerMet 36

Qy 277 CTGACGACGACGAGTCTTACAG-----GTCTGACGACGACGACGACGACGACG 327

Db 37 ILeGluHisAspSerTyrTyrLysAspGlnSerHisLeuThrPheGluGluArgValLys 56

Qy 328 GCCTTGAAGACAGTACATTTTGACCATCCAGATCGCTTGTATTAATGATTGATGAC 387

Db 57 Thr-----AsnTyrAspHisPheProAlaPheAspThrAspLeuMetIle 71

Qy 388 AGGACTGTGAAGACATCGTGTGAGGCAAAACGTTGAGGCGGACCTTATGATTGTG 447

Db 72 GluGlnIleLysGluLeuAlaGlyLysArgProValAspIleProThrTyrAspTyrThr 91

Qy 448 ACACACTCAAGTTCACGACGACGACGACGACGACGACGACGACGACGACGACG 507

Db 92 GluHisThrArgSerSerLysThrTyrArgGlnLysProGlnAspValPheIleValGlu 111

Qy 508 GGCATCTTGTGTCTACAGCAGGACGACGACGACGACGACGACGACGACGACG 567

Db 112 GlyIleLeuValLeuGluAspLysArgLysArgLysArgLysArgLysArgLys 131

Qy 568 GACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 624

Db 132 AspThrAspAspAspValArgIleLysArgIleLysArgIleLysArgIleLysArg 151

Qy 625 AGGACCTGAGCAGATCTGACACACACCTTGTGAAGCCGCGCTTGAGAG 684

Db 152 ArgSerLeuAspSerValIleAsnGlnTyrLeuGluValValLysProMetTyrHisGln 171

Qy 685 TTCTGCTCGGACCAAGAGATGCGGATGATCATCCACGAGAGTGAATATG 744

Db 172 PheIleGluSerThrLysArgTyrAlaAspIleValIleProGluGlyValSerAsnThr 191

Qy 745 GTTGCCATCAACCTGATCGTGACACATCCAGACATCTTG 786

Db 192 ValAlaIleAspLeuThrThrLysIleAlaLysIleLeu 205

RESULT 10

S29374

uridine kinase (EC 2.7.1.48) - yeast (Saccharomyces cerevisiae)

N/Alternate names: protein N2050; protein YNR012w

C/Species: Saccharomyces cerevisiae

C/Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 23-Mar-2001

C/Accession: S29374; S45134; S48349; S63338; S21361

R/Kern, L.

Nucleic Acids Res. 18, 5279, 1990

A/Title: The URK1 gene of Saccharomyces cerevisiae encoding uridine kinase.

A/Reference number: S29373; MUID:90384830; PMID:2169608

A/Accession: S29374

A/Molecule type: DNA

A/Residues: 1-501 <KOR>

A/Cross-references: EMBL:X53998; NID:G4771; PIDN:CAA37946.1; PID:G4773

R/Vernasselt, P.; Aert, R.; Voet, M.; Voelckaert, G.

submitted to the EMBL Data Library, January 1994

A/Description: Twelve open reading frames revealed on the 23.6 kbp segment flanking the

A/Reference number: S45119

A/Accession: S45134

A/Molecule type: DNA

A/Residues: 1-501 <VER>

A/Cross-references: EMBL:X77395; NID:G496717; PID:G496728

R/Vernasselt, P.; Aert, R.; Voet, M.; Voelckaert, G.

Yeast 10, 1355-1361, 1994

A/Title: Twelve open reading frames revealed in the 23.6 kb segment flanking the centromere

A/Reference number: S48338; MUID:95208356; PMID:7900425

A/Accession: S48349

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-501 <VE2>

A/Cross-references: EMBL:X77395; NID:G496717; PIDN:CAA54580.1; PID:G496728

R/Aert, R.; Vernasselt, P.; Voet, M.; Voelckaert, G.

submitted to the Protein Sequence Database, April 1996

A/Reference number: S62910

A/Accession: S63338

A/Molecule type: DNA

A/Residues: 1-501 <AER>

A/Cross-references: EMBL:271627; NID:G1302489; PID:2339785; PID:G1302490; MIPS:YNR012w

A/Experimental source: strain S288C

C/Genetics:

A:Gene: SGD:URK1

A/Cross-references: SGD:S0005295; MIPS:YNR012w

A/Map position: 14R

C/Keywords: phosphotransferase; transmembrane protein

F/154-170/Domain: transmembrane #status predicted <TM1>

F/419-435/Domain: transmembrane #status predicted <TM2>

Alignment Scores:

Pred. No.:	2.1e-23	Length:	501
Score:	437.50	Matches:	97
Percent Similarity:	53.76%	Conservative:	46
Best Local Similarity:	36.47%	Mismatches:	88
Query Match:	14.75%	Indels:	35
DB:	2	Gaps:	5

US-09-896-522-1 (1-1624) x S29374 (1-501)

Qy 85 GAGGCCGAGATGCTTCGCGGAGGCGGAAGCTCCGAGAGCCCGCGGAGCCGAC 144

Db 28 LysAlaAsnAlaValMetAspGlyGluValAspValLysLysThrLysLysSer 47

Qy 145 CGT-----CCGACACAGCGGCGCTTCTGATGAGGCGTGAAGCGGACCTGCCACC 195

```

Db      48 ArgTyrIleProProtrpThrThrProTyrIleIleGlyIleGlyAlaIleSerylSer 67
QY      196 GGAAGTCCAGCCCTGTCTGTGAAGATCATGAGTGTCTGGGACAGAACGAGTGGAAACAG 255
      68 GilylserThservalAlaAlaIleValSerSertle----- 80
QY      256 CGGACGCGAAGGTGGTCATCTGAGCCAGCAGGTTCTTACAAAGGTCCTGACGGCAGAG 315
      81 AsnValProTrrpThrValIleuIleSerLeuAspAsnPhetYrAsnProIleuGlyProGlu 100
QY      316 CAGAGCGCCAGGCGCTTGAAGAAGACAGTACATTTTTCAGCATCCAGATGCCCTTGTATAT 375
      101 AspArgAlaIleArgAlaPhetYrAsnGluTyrAspPhetSpgIupProAsnAlaIleAsnLeu 120
QY      376 GATTGTATGACACGAGCTGTGAAGAACATCTGGAGGGCCAAACGGTGGAGGCCGAC 435
      121 AspleuAlaIleTyrIleCysIleAsnLeuIleGluIleGlyIleTyrIleAsnIleProVal 140
QY      436 TATGATTTTGTGACACACTCAGGTTACAGAG---ACCAAGGTGGTCTACCTCGGCGAC 492
      141 TyrSerPheValIleHisAsnArgValProAspIleAsnIleValIleTyrGlyAlaSer 160
QY      493 GTGGTCTGTGTGAGGCGCATCTTGTTGTCTACAGCCAGAGATCCGAGCATGTTCCAC 552
      161 ValValValIleGluIleGlyIleTyrAlaLeuTyrAspArgIleuLeuAspIleuMetAsp 180
QY      553 CTGCGCCTCTTCTGTGACACCGACTCCGAGTCTGAGGTCGTGGAAGTTCCTCGGAC 612
      181 LeuIleValIleTyrValAspAlaAspLeuAspValCysLeuAlaArgIleuSerArgAsp 200
QY      613 ---GTGGCGCGAGGAGGAGGAGCTGAGCAGATTTCTGACAGTACACACCTTCTGTAAG 669
      201 IleValSerArgIleGlyAspLeuAspIleCysIleGlnIleTrrpGlyIlePheValIle 220
QY      670 CCGGACCTTCAGAGATCTCTGCTCCGACAAAGATATGCCATGTGATCATCCACGA 729
      221 ProAsnAlaValIlePheValIlePheValIlePheValIlePheValIleIlePheSer 240
QY      730 GAGGTGACAAATATGTTGCCATCAACCTGATCGTCGACGACATC----- 774
      241 MetSerAspAsnAlaThrAlaValAsnLeuIleIleAsnHisIleIleSerylSleuGlu 260
QY      775 -----CAGAG 780
      261 LeuIleSerAsnGluHisLeuArgGluLeuIleIleSerylSerProSerGlnAsp 280
QY      781 ATTCTGATGTGATCATC 798
      281 ValIleuAsnArgAsnIle 286

```

RESULT 11

G69728
uridine kinase udk - Bacillus subtilis
C/Species: Bacillus subtilis

C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000

C/Accession: G69728
R/Kunze, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A.; Ehrlich, S.D.; Emerson, P.T.; Ertian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizi, A.; Gallier

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.P.

Koecker, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, A.; Lapidus, A.; Lardinois,

A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli

Rieger, M.; Rivolta, C.; Roche, B.; Roche, M.; Sadale, Y.; Sato, T.; Scanlon,

A/Authors: Schleicher, S.; Schroeter, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Serot

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Toosto, V.; Uchiyama,

T.; Winters, P.; Wipet, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A/Authors: Yoshikawa, H.F.; Zunsstein, E.; Yoshikawa, H.; Danchin, A.

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A/Reference number: A69580; MUID:98044033; PMID:9384377

A/Accession: G69728

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-211 <R>

A/Cross-references: GB:299117, GB:AL009126; NID:g2634966; PIDN:CAB14675.1; PID:g2635179

A/Experimental source: strain 168

C/Genetics:

A/Gene: udk

C/Superfamily: uridine kinase

Alignment Scores:

Pred. No.:	Length:	211
Score:	6,39e-23	84
Percent Similarity:	430.00	52
Best Local Similarity:	63.26	63
Query Match:	39.07	16
	14.49	4
	Gaps:	

DB: 2

US-09-896-522-1 (1-1624) x G69728 (1-211)

```

QY      154 CAGCGGCCCTTCTGATAGGGGTGAGCGGCGGACGCGGAGGAGTGCAGCGGTGT 213
      3 LysAsnProValValIleGlyIleAlaGlyIleGlySerGlySerGlyIleYrIleThrValThr 22
QY      214 GAGAGATGATGAGGTGCTGGGACAGAAAGAGTGAACAGCGGAGCGGAGGAGTGT 273
      23 ArgSerIleTyrGlu-----GlnPheIleGlyIleSerIleIleLeu 35
QY      274 ATCTTGAGCCAGGACAGGTTCTACAG-----GTCTGACGCGGAGGAGGAGGCC 324
      36 MetIleGlnGlnAspIleuTyrIleYrAspGlnSerIleuIleuProIleGluIleuIleu 55
QY      325 AAGCGCTTGAAGGAGAGTGAACATTTTGAACATTCAGATGCCATGATGATGATGATG 384
      56 AsnThr-----AsnTyrAspHisIleuIleuIleuIleuIleuIleu 70
QY      385 CACAGAGCTCTGAAGAACATCTGTGAGGCGGAAACGCGGAGGAGTGCAGCTATGATTT 444
      71 IleGlnHisIleGlnAspIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 90
QY      445 GTGACACACTCAAGGTTACCGAGACGAGTGTCTACCTCGGAGCGGAGTGTGTT 504
      91 LysLeuHisIleThrArgSerGluIleuIleuIleuIleuIleuIleuIleuIleu 110
QY      505 GAGGAGATCTGTGTGTCTTACAGCCAGGAGATCCGCGGAGATGTCACCTCGGCGCTTC 564
      111 GlnGlyIleLeuValIleuGluAspIleuIleuIleuIleuIleuIleuIleuIleu 130
QY      565 GTGACACGAGCTCGAGGTCAGGCTGTCTCGAAGAGTTCCTCGGAGCGT---CGCGGA 621
      131 ValAspThrAspAlaAspLeuArgIleIleIleIleIleIleIleIleIleIleIleIle 150
QY      622 GGGAGGAGCTTGAGAGATTTCTGACGACATCACACCACTTCGTGAACCGGCGCTTCGAG 681
      151 GlyArgSerIleAspIleuValIleGluIleuIleuIleuIleuIleuIleuIleuIleu 170
QY      682 GAGTTTGTCTGCGGAGGAGGAGTATGCGAGTATCATCCAGAGAGAGTGAAT 741
      171 GlnPheValIleuProThrIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 190
QY      742 ATGGTTCATCACTGATCGTGTGAGCAGCATCATCCAGATGAGCATTTGTG 786
      191 HisValAlaIleAspIleuMetValIleuIleuIleuIleuIleuIleuIleuIleu 205

```

RESULT 12

uridine kinase [imported] - Halobacterium sp. NRC-1
F84277

C/Species: Halobacterium sp. NRC-1
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002

C/Accession: F84277
R/Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laeky, S

R.; Leithauer, B.; Koller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor

Jung, K.H.; Alam, M.; Frettas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of *Halobacterium* species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: F84277
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-248 <STO>
A:Cross-references: GB:AE004437; NID:g10580750; PIDN:AA619586.1; GSPDB:GN00138
C:Genetics:
A:Gene: urk
C:Superfamily: uridine kinase

Alignment Scores:
Pred. No.: 5.58e-22 Length: 248
Score: 417.00 Matches: 83
Percent Similarity: 63.64% Conservative: 50
Best Local Similarity: 39.71% Mismatches: 58
Query Match: 14.05% Indels: 18
DB: 2 Gaps: 5

US-09-896-522-1 (1-1624) x F84277 (1-248)

```
Oy 163 TTCCGTAGGGGAGCGGCGGCACTGCCAGCGGAACTGCGTGTGAGAGATC 222
Db 29 PhehAlleglylleAlaGlyGlyThrclyAlaGlysthrtrValAlaArgGluLe 48
Oy 223 ATGAGTGTCTGGGACAGACAGAGGTGGACAGCGGCGGAGAGTGTGATCCTGAGC 282
Db 49 ThrAspAsnValGlyGluSer-----AlaThrLeuIlePro 60
Oy 283 CAGACAGGTTCTTCAAGAGCTCTGACG-----GCAGACGAAGGCCAGGCTTG 333
Db 61 LeuAspAsnTrpTrpGluAspLeuSerAspArgProPheGluGluArgAlaAsnAla--- 79
Oy 334 AAAGACAGATCAATTTTGAACATGCCAGATGCCCTTGAATATGATTTGATGACAGACT 393
Db 80 -----AsnTrpAspHisProSerAlaPheGluTrpGluLeuLeuArgThrHis 95
Oy 394 CTGAAGAACAATCTGTGAGAGCGCAAAACGATGAGGTGGCGACCTATGATTTTGTGACAC 453
Db 96 IleAspAlaLeuLeuSerGlyGlnSerIleGluMetProGlnTrpAspPheGluArgHis 115
Oy 454 TCAGAGTTTACAGACACCGGTGCTTACCTTCGCGACGCTGCTGTTGAGGGCATC 513
Db 116 ValArgGlyAlaAspArgValValValGluProThrAspValIleValIleGluGlyIle 135
Oy 514 TTGGTGTTCACACCGAGATCCGGGACATGTTCCACCTCGCGCCCTTCCTGAGACAC 573
Db 136 LeuAlaLeuSerAspGluThrValAspAspMetLeuAspLeuHisIleTrpValGluThr 155
Oy 574 GACTCCGACGTCAGCTGTCTCGAAGAGTTCTCCGGAC--GTGCCCGCAGAGGAGGAC 630
Db 156 AspAlaAspValAlaGlyIleLeuArgArgIleGluArgAspValValGluArgGlyArgGlu 175
Oy 631 CTGAGACAGATTTGACGCGCATACCACTTCTGAAGCCGGGCTTCGAGAGTTTGC 690
Db 176 LeuGluGlyValMetAspArgTrpLeuSerThrValIlyspromethisGluGlnPheIle 195
Oy 691 CTGGCCGCAAGAAGATGCGCATGATCCACAGAGAGATGAGCAATATGCTTGGC 750
Db 196 GluProThrTrpAspHisAlaAspIleIleProGluGlyAla--AsnSerValAla 214
Oy 751 ATCAACTGATCTGTCAGACATCCAG 777
Db 215 ValAsnLeuLeuGluGluValGln 223
```

RESULT 13

D86582
uridine kinase [imported] - *Chlamydomonas reinhardtii* (strain J138)
C:Species: *Chlamydomonas reinhardtii*, *Chlamydomonas reinhardtii*
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: D86582
R/Shitai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is

Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of *Chlamydia pneumoniae* J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: D86582
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-222 <STO>
A:Cross-references: GB:BA000008; NID:g8979107; PIDN:BA98942.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: CP10735
C:Superfamily: uridine kinase

Alignment Scores:
Pred. No.: 1.25e-21 Length: 222
Score: 412.00 Matches: 84
Percent Similarity: 63.01% Conservative: 54
Best Local Similarity: 38.36% Mismatches: 59
Query Match: 13.89% Indels: 22
DB: 2 Gaps: 6

US-09-896-522-1 (1-1624) x D86582 (1-222)

```
Oy 166 CTGATAGGGGTGACGCGGCGGCACTGCCAGCGGAACTGCGTGTGAGAGATCAG 225
Db 8 IleIleGlyIleThrGlySerGlyAlaGlyLeuThrThrLeuThrGlnAsnIleLys 27
Oy 226 GAGTGTCTGGGACAGACAGAGGTGGACAGCGGCGGAGAGTGTGATCCTGAGCCAG 285
Db 28 GluIlePheGlyGluAsp-----ValSerValIleCysGln 39
Oy 286 GACAGGTTCTTCAAG-----GTCTGACGCGCAGACAGAGGCCCAAGGCTTGAA 336
Db 40 AspAsnTrpTrpTrpAspArgSerHisTrpThrProGluGluArgAlaAsnLeuIle---- 58
Oy 337 GGAAGTCAATTTTGAACATGCCAGATGCCCTTGAATATGATTTGATGACAGACTTG 396
Db 59 -----TrpAspHisProAspAlaPheAspAsnAspLeuIleSerAspIle 74
Oy 397 AAGAATGCTGTGAGGCGCAAAACGGTGAGAGTGGCGACCTATGATTTTGTGACACTCA 456
Db 75 LysArgLeuLeuAsnAsnGlnIleValGlnAlaProValPheAspPheValIleGluGln 94
Oy 457 AGG---TTACAGAGACCAACGCGGTCTACCTTCGCGAGAGTGTGTTGAGGCGATC 513
Db 95 ArgSerLysThrGluIleGluThrIleTrpProSerLysValIleLeuValGluGlyIle 114
Oy 514 TTGGTGTTCACACCGAGATCCGGGACATGTTCCACCTGCGCCCTTCCTGAGACAC 573
Db 115 LeuValPheGluAsnGlnGluIleLeuArgAspLeuMetAspIleArgIlePheValAlaPthr 134
Oy 574 GACTCCGACGTCAGCTGTCTCGAAGAGTTCTCCGGAC--GTGCCCGCAGAGGAGGAC 630
Db 135 AspAlaAspGluArgIleLeuArgArgMetValArgAspValGlnIleGluGlnAspSer 154
Oy 631 CTGAGACAGATTTGACGCGCATACCACTTCTGAAGCCGGGCTTCGAGAGTTTGC 690
Db 155 ValAspCysIleMetSerArgTrpLeuSerMetValIlyspromethisGluGlnPheIle 174
Oy 691 CTGGCCGCAAGAAGATGCGCATGATCCACAGAGAGATGAGCAATATGCTTGGC 750
Db 175 GluProThrArgLysTrpAlaAspIleIleValIleGlyAsnTrpArgGlnAsnValVal 194
Oy 751 ATCAACTGATCTGTCAGACATCCAG 777
Db 195 ThrAsnIleLeuSerGlnLysIleLysAsnHisLeuGluAsnAlaLeuGluIleSerAsp 213
```

RESULT 14

E72041
uridine kinase CP0011 [imported] - *Chlamydomonas reinhardtii* (strains CWL029 and AR39)
C:Species: *Chlamydomonas reinhardtii*, *Chlamydomonas reinhardtii*
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: E72041; E81622

R:Kaltman, S.; Hitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Ollinger, L.; Grimwood, J.
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: E72041
A:Molecule type: DNA
A:Residues: 1-222 <RNA>
A:Cross-references: GB:AE001655, GB:AE001363; NID:g4377039; PIDN:AA018874.1; PID:g4377070
A:Experimental source: strain CML029
R:Read, T.D.; Brumham, R.C.; Shen, C.; Gill, S.R.; Heidelberger, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwin, J.M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1405, 2000
A:Title: Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39.
A:Reference number: AB1500; MUID:20150255; PMID:10684935
A:Accession: B81622
A:Molecule type: DNA
A:Residues: 1-222 <REA>
A:Cross-references: GB:AE002165; GB:AE002161; NID:g7188948; PIDN:AA937907.1; PID:g7188929
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CPN0735, CP0011
;Superfamily: uridine kinase

Alignment Scores:	
Pred. No.:	1-25e-21
Score:	412.00
Percent Similarity:	63.01%
Best Local Similarity:	38.36%
Query Matchn:	13-89%
2	
Length:	222
Matches:	84
Conservative:	54
Mismatches:	59
Indels:	22
Gaps:	6

[illegible]

226 GAGTTGCTGGACAGAACGAGTGGAA CAGCGGACGGGAGGTGTCTATCTGAGCCAG 285
 |||:::|
 28 GtutlPhehelglycluasP-----ValserValIleCygsIn 39
 |||:::|

[illegible]

QY 337 GGACGATACATTGTGACCATCCAGATGCCCTTGTATATGATTGATGCACAGACTCTG 396

bb 59 -----TTPASbhisPProASpAlaPneASpASnSpLeuLeuIesASnIle 74

Cy 397 AAGAACTCGTAGGAGGC AAAACGGTGAAGTGCCACCATATGATTTTGTGACACACTCA 456
:: :: :: :: :: :: :: :: :: :: :: ::
Db 75 LysATcLeuIvsAAsnGluIleValIGlnAlaPheValIbeaDnsPheValIengIvsAn 94
:: :: :: :: :: :: :: :: :: :: :: ::

457 AGG---TTACGAGACACCGAGGTCTACCCGTGGAGCGTGTTCTGTTGAGGCATC 513
OY |||||
DB ArgsetIwmbhrGjlltIcglumtrllctIyProsetIwvjaIteIevvaIgluGjlltIe 114
95 |||||

Oy
514 TTGATCTTACAGCCAGAGATCCGGACAATTGCCACCCTTCGTGGACACC 573

Dh
115 [LWV]PheGLI[ENG][C]LeuThrHsdRLeuMetAsnTrpHisPheValAlaThr 134

0y 574 GACTCGAGCTGCAGGCTGTCTCGAAGATTCCGGGACGTGCC---CGAGGGAGGGAC 630
|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db 135 AAAATAATTCCTTTTAACTCCTTAATCTTGCTGAAGAAGCAAGAAC 154

[illegible]

691 CTGCCGACAAAGAGTATGCCGATGTGATCATCCACGAGAGTGGACATATGCTGCC 750

QY	751	ATCAACCGTATCGGACG-----CACATCCAGAGACATTTCGAATGTGAC	795
Db	195	ThrsnllleueuerglnlyslleLysasnhlslneuglualsnlalebgluserap	213

RESULT 15
G70101
uridine kinase (udk) homolog - Lyme disease spirochete

CiDate: 13-Feb-1998 #sequence_revision 13-Feb-1998 #next_change 17-Nov-2000
#RFAccession: G70101, 146979
RFAcraer, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White

A. Authors: Smith, H.O.; Venter, J.C.
C. Editors: Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
N. Nature 390, 580-586, 1997

[illegible]

A: Molecular type: LMA
A: Residues: 1-207 <KLE>
A: Cross-references: GB:AE001115; GB:AE000783; NID:G2687879; PIDN:AAC6392.1; PID:G2687886
A: Experimental source: strain B31

A. Reference number:224339. PMID:97372541. PMID:9228761

A:Accession: 146979
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 2-207 <BOU>

A:Cross-references: EMBL:X57449; NID:g1405440; PIDN:CNA66081.1; PID:g1405442
A:Experimental source: strain HB19, ssp. sensu stricto
C:Genetics:
A:Gene: uidK

C:Superfamily: uridine kinase	
Alignment Scores:	
Seed No.: 1	59e-21
Length:	207

Score:	410.50	Matches:	79
Percent Similarity:	59.81%	Conservative:	49
Best Local Similarity:	36.92%	Mismatches:	59
Query: March	13.84%	Indels:	27

DB: 2 Gaps: 4
US-09-896-522-1 (1-1624) x G70101 (1-207)

QY CTGAATGAGGGGTGACCGCGGCACCTGCCAGCCGAGTCAACCCGTGTGAGAAGATCATG 225
166 :::::
Db 5 TTTCTGGTATTCerGlySerGlySerCylLysThrValValSerLysIleSer 24

QY	226	GAGTTGCTGAGGAAGAA	CGAGGTGAA	CACGCGACGCGAAGG	GGTCATCCTGAGCAG	285
		:::		:::	:::	
db	25	GluphelleProcin	-----	RhevalleutiesergIn		35

QY 286 GACAGTTCCTACAGCTCTACGGCAGACAGAAAGGCCAAGGCTTGAAGACACTAC 345
 Db 36 AspAsnTyrTyrIysSerVal-----GlyAspTyr 45

QY 346 -----AATTTGACCATCCAGAGCCCTTTGATTAATGATTGATG 384

Db 46 GIuHISGIuPheSerLysValAsnIleAspHisIleProAspAlaPheAspAsnIleLeuPhe 65

QY 385 CACAGGACTCTGAAGACATCTGTGGAGGGCAAAAACGGTGGAGGTCCGACCATATGATTTT 444

Db 66 TGTGTAHISLEULYASNLLEULYLYSASNSERPROILEASPHETPROLEUITYAAPH 85

QY 445 GTGACACATCTAAGGTTTACCGAGACACCGGTGCTCTACCTCGGACGTGGTCTCTTT 504
Db 86 TleantHslsyrgrglnleuylstrvalleuvalvalProthrProvalValleval 105

QY 505 GAGGCACTTGTGTTCTACAGCCAGAGAATCCGGACATGTCACCTGCCTCTTC 564
||||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

[illegible]

```
Db      106 GluGlyIleMetIlePheValGluGluArgValArgAsnLeuIleAspLeuLysIleTyr 125
QY      565 GTGGACACCGACTCGACTCAGGCTGTCTCGAAGATTCTCCGGGACGTG--CGCCGA 621
      :::||||| :::||||| :::||||| :::||||| :::|||||
Db      126 IleAspThrProAsnAspIleArgPheIleArgIleuArgArgAspIleSerLysArg 145
QY      622 GGGAGGGACCTGAGACAGATTCTGACGCGAGTACACACCTTCGTGAAGCCGGCCTTCGAG 681
      ||||| :::||| :::||| :::|||
Db      146 GlyArgThrValGluSerValIleAspGlnTyrLeuAsnThrThrArgTrpGlyTyrTyr 165
QY      682 GAGTTCTGCCCTCCGACAAAGATATGCCGATGTGATCATCCACGAGAGAGTGACAAAT 741
      ||||| :::||| :::||| :::||| :::|||
Db      166 ArgPheIleGluProThrLysGlnTyrAlaAspLeuIleIleProGluGlyGlyHisAsn 185
QY      742 ATGGTTGCCATCAACTGATCGTCGACGACATCCAGACATT 783
      ||||| :::||| :::||| :::|||
Db      186 AspLysAlaLeuTyrValLeuSerThrPheLeuLysSerLeu 199
```

Search completed: November 25, 2003, 07:40:36
Job time : 48.9992 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model1

Run on: November 25, 2003, 07:30:54 ; Search time 153.943 Seconds
(without alignments)
3891.516 Million cell updates/sec

Title: US-09-896-522-1

Perfect score: 2967
Sequence: 1 gggggggtcgccctccgagctc.....ccagctgtgtgggggagcag 1624

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 673684 seqs, 184443283 residues

Total number of hits satisfying chosen parameters: 1347368

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09896522/runcat_21112003_184107_2917/app_query.fasta_1.2830
-DB=Published Applications_AA -QFMT=fastcat -SUFFIX=n2p.rapb -MIMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS-human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=acct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEARSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09896522 @CGN_1_1_129 @runcat_21112003_184107_2917
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubppa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubppa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubppa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubppa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	1450	48.9	277	9	US-09-896-522-2
2	925	31.2	337	10	US-09-925-300-1160
3	428.5	14.4	125	9	US-09-896-522-5
4	327	11.0	60	12	US-10-029-386-28765
5	320.5	10.8	124	9	US-09-896-522-6
6	276	9.3	231	9	US-09-896-522-4
7	192.5	6.5	797	15	US-10-156-761-10907
8	177	6.0	595	10	US-09-738-973-187
9	177	6.0	595	10	US-09-854-133-187
10	177	6.0	595	10	US-10-144-649A-187
11	172.5	5.9	503	14	US-10-078-547-2
12	169.5	5.8	265	11	US-09-975-719-369
13	169	5.7	1207	15	US-10-024-368-5
14	164	5.6	507	14	US-10-078-547-24
15	164	5.6	737	12	US-10-017-161-1350
16	163	5.5	1535	15	US-10-189-971-14
17	163	5.5	1570	15	US-10-189-971-12
18	162	5.5	1182	15	US-10-024-368-6
19	159.5	5.4	1063	12	US-10-017-161-1624
20	158.5	5.4	375	15	US-10-156-761-14622
21	158.5	5.4	1593	15	US-10-189-971-4
22	158.5	5.4	1628	15	US-10-189-971-2
23	156.5	5.3	877	12	US-10-237-496-90
24	156.5	5.3	877	12	US-10-242-074-90
25	156.5	5.3	877	12	US-10-242-505-90
26	156.5	5.3	877	12	US-10-242-574-90
27	156.5	5.3	877	12	US-10-242-261-90
28	156.5	5.3	877	12	US-10-243-282-90
29	156.5	5.3	877	12	US-10-243-402-90
30	156.5	5.3	877	12	US-10-243-431-90
31	156.5	5.3	877	12	US-10-244-164-90
32	156.5	5.3	877	12	US-10-244-972-90
33	156.5	5.3	877	12	US-10-197-942-90
34	156.5	5.3	877	12	US-10-238-136-90
35	156.5	5.3	877	12	US-10-245-013-90
36	156.5	5.3	877	14	US-10-041-770-2
37	156.5	5.3	877	15	US-10-245-103-90
38	156.5	5.3	877	15	US-10-245-107-90
39	156.5	5.3	877	15	US-10-245-113-90
40	156.5	5.3	877	15	US-10-245-771-90
41	156.5	5.3	877	15	US-10-245-883-90
42	156.5	5.3	877	15	US-10-245-851-90
43	156.5	5.3	877	15	US-10-237-535-90
44	156.5	5.3	877	15	US-10-238-183-90
45	156.5	5.3	877	15	US-10-238-283-90

ALIGNMENTS

RESULT 1
US-09-896-522-2
; Sequence 2, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-896-522-2

Alignment Scores: 4.26e-101 Length: 277
Pred. No.: 1450.00 Matches: 277
Score:

Percent Similarity:	100.00%	Conservative:	0
Best local Similarity:	100.00%	Mismatches:	0
Query Match:	48.87%	Indels:	0
DB:	9	Gaps:	0

US-09-896-522-1 (1-1624) X US-09-896-522-2 (1-277)

OY		9	ATGGCTTCGGCGGAGGAGCCGAAGCACTCGAGAGCGCCGCGCGGAGGCGGACCTGCGCAC	153
Db		1	MetAlaSerMetActGlyGlyAlaMetArgCysGlnSerProAlaProGluAlaMetArgProHis	20
OY		154	CAGCGGCCCTTCTGTATAGGGAGGTGAGCGGCGGCGACTGCCAGCGGAGTGCACCTGTGT	213
Db		21	GlnArgProPhePheLeuIleGlyAlaSerGlyGlyThrAlaSerGlyLysSerThrValLys	40
OY		214	GAGAAAGTCAATGAGAGTTGCTGGGACAGCAACGAGGTGGAAACGCCGACACGGGAAGTGTGTC	273
Db		41	GlnLysLeuMetGlnLeuLeuGlnGlyGlnAsnGlnValGlnGlnAlaArgGlnArgLysValVal	60
OY		274	ATCCTGAGCCAGGACAGAGTTCTTCAAGAGTCTCGAGGCGAGACAGACAAAGCCAAAGCCTTG	333
Db		61	IleLeuSerGlnMetArgPheThrArgPheTyrLysValLeuThrAlaGlnGlnLysAlaLysAlaLeu	80
OY		334	AAAGGACAGTCAAAATTTTGAACATCCAGATGCCCTTGAATGATTTGATGCACAGACT	393
Db		81	LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAspMetLeuMetHisArgThr	100
OY		394	CTGAAGAAACATCGTGGAGGGCAAAACGGTGGAGGAGTGGCCGACCTGATGATTTGTGACAC	453
Db		101	LeuLysAsnLysLeuAlaGlnGlyLysThrValGlnValProThrTyrAspPheValThrHis	120
OY		454	TCAAGGTTACCAAGACCAACGAGTGTCTACCTCGGAGCTGGTTCGTGTTGAAGGCATC	513
Db		121	SerArgPheProGlnThrValThrValValTyrProAlaAspValValLeuPheGlnGlyIle	140
OY		514	TTGGTGTTTACACGACGAGAGATCGGGGACATGTTCCACTGGCGCTCTTGCTGGGACAC	573
Db		141	LeuValaPheTyrSerGlnGlnLysLeuArgPheMetPheHisLysLeuArgLeuPheValAlaSprThr	160
OY		574	GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGAGCTGCGCGGACGCGGAGGAGGACCTG	633
Db		161	AspSerMetArgValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgArgAspLeu	180
OY		634	GAGCAGATTTGTAGCGAGTACACACCACTTCGTAAGCCGGGCTTGAGAGATTTGTGCTG	693
Db		181	GlnGlnLysLeuThrGlnTyrThrThrPheValLysProAlaPheGlnGlnPheCysLeu	200
OY		694	CCGACAAAGAGTATGCGCATGTGATCATCCACAGAGAGTGGACATATGATTGTTGCCATC	753
Db		201	ProThrLysLysTyrAlaMetArgValIleIleProArgGlyValaAspHisMetValaIle	220
OY		754	AACCTGATCGTGACGACATCCAGGACATTTGAAATGTTGACATGTGCAAAATGACACCGA	813
Db		221	AsnLeuLysValGlnHisIleGlnMetPheIleLeuAsnGlyAspLysLysTyrPheArg	240
OY		814	GAGAGGTCCTCAATGGGGGAGCTACAGCGGACCTTTCTGAGGCCAGGGGACACACCTGGG	873
Db		241	GlyGlySerAsnMetLysArgSerTyrLysArgThrPheSerGlnProGlyAspHisArgProGly	260
OY		874	ATGCTGACCTCTGGCAACCGGTGCATTTGAGTCCAGACGACGACCCAC	924
Db		261	MetLeuThrSerGlyLysArgSerHisLysGlnSerSerArgProHis	277
RESULT 2				
US-09-925-300-1160				
; Sequence 1160, Application US/09925300				
; Patent No. US20020151681A1				
; GENERAL INFORMATION:				
; APPLICANT: Craig Rosen,				
; APPLICANT: Steve Ruben				
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies				
; FILE REFERENCE: PA101				
; CURRENT APPLICATION NUMBER: US/09/925,300				

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

US-09-896-522-1 (1-1624) x US-09-925-300-1160 (1-337)

QY 12 TCCGACCTCGGCGCTGGGGGGGGCGGGGCCCGGAGGCGGGCGGAGCCG 71

Db 54 SerAspAlaIeYlaIeY---GlyGluArgAlaSerValArgThrGlySerGly----- 70

QY 72 ATCGCGGGAGCGGAGCGCCGAGATGGCTTCGGCGGGAGCGGAGAACTGCGAGACCCCG 131

Db 71 ---ArgArgGlyGlyAlaAsnHisIeGlyArgGlyGlnArgAlaAspProAlaGluProPro 89

QY 132 GCCGAGGCGGACCGTCGCGCACCGAGCGCCTTCCTGATAGAGGGGTGAGCGGCGACTGC 190

Db 90 AlaAlaGlnArg-----ArgArgAlaLeuProTyrArgArgHisGlyGlyThrAl 106

QY 192 CAGCGGGAAAGTCGACCGCTGTGTGAGAAAGATCATGAGATTGCTGGGACAGAAACGAGTGA 251

Db 106 aserGlyYlserSerValCysAlaIaYsIeValGlnIleuLeuGlyGlnIasnGluValAs 126

QY 252 ACAGCGGACGCGGAGAGGTGATCATCTGAGCGAGGACAGGTTCTCAACAGCGCTCGAGCGC 311

Db 126 pTyrArgGlnYsGlnValValIleIleuSerGlnIasnSerPheTyrArgValIleuTrse 146

QY 312 AGAGCGAAAGGCCAGAGCCTTGAAAGGACAGTACATTTGACCATTCAGATGCGCTTTGA 371

Db 146 rGlnGlnYsAlaYsAlaYsAlaIeLeuYs**GlnPheIasnPheIserProAspAlaPheAs 166

QY 372 TAAATGATTGATGCACAGGACTCTTAAGAAACATCTGGAGGGCAAAACGGTGAAGTGCC 431

Db 166 pAsnGln***IleLeuYsThrLeuYsGlnIleThrGlnGlyYsThrValGlnIlePr 186

QY 432 GACCTATGATTTTGTGACACTCAAGGTTCACGAGACACAGGAGGTCTTACCTGTGGGA 491

Db 186 oValTyrAspPheValSerHisSerArgYsGlnIuThrValThrValTyrProAlaAs 206

QY 492 CGAGGTTCGTTTGAAGGCGATCTTGATGTCTTCAACGCGAGAGATCCGGGACATGTCCA 551

Db 206 pValIaIleuPheGlnGlyIleLeuAlaPheTyrIserGlnIuValIArgAspIleuPheGln 226

QY 552 CTGCGGCTCTTCGTGAGACCGAGCTCGACGTCAGCTGTCTCGAAGAGTTCTCCGGGA 611

```

Db      226  nmtelysleuphevalasptrhaspalaspthrargyleuserargvalleuhargas
               ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Qy      612  CGTG---CGCCGAGGAGGAGCACTTGAGCAGATTTCAGCGCAGTACACCACTTCGTGAA
               ::::|
Db      246  pleserlugarglyargaspheulugnilleusergintlyrillehrhethally
               |||
Qy      669  GCCGCGCTTCGAGAGAGTTTCGCTGCCGCAAGAAAGATATGCCAGTGTGATCATCCACG
               |||
Db      266  srolaPheglugluPheCysleuProthrlyslstyAlaaspvallelleproar
               |||
Qy      729  AGGAGTGACAAATATGGTGGCATCAACCTGATCGTGAGCAGCATCCAGCACTTCGTGA
               |||
Db      266  golyalaspashevalalallemsneuilevalghnhsilleghnspilleuas
               |||
Qy      789  TGGTGACATCTGCMAATGCGACCCGAGAGGGTGTCCAATGGCGGAGCTTACAGCGGACCTT
               |||
Db      306  ngly-----GlyProserlylsargdlnthr--
               |||
Qy      849  TTCTGAGCCAGGGGACCACTTCGGGATGCTGACCTTCGGCAACGGTCACATTGGAGTC
               |||
Db      315  -----AsnglyCysleuasnglyThrProserArglylsArglnalasergluse
               |||
Qy      909  CAGCAGCAGACCCCCAC 924
               |||
Db      332  rsererargProhis 337
               |||

RESULT 3
US-09-896-522-5
: Sequence 5, Application US/09896522
: Patent No. US20020055161A1
: GENERAL INFORMATION:
: APPLICANT: Glucksmann, Maria A.
: TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
: TITLE OF INVENTION: USES THEREOF
: FILE REFERENCE: 381552001700
: CURRENT APPLICATION NUMBER: US/09/896,522
: CURRENT FILING DATE: 2001-06-28
: PRIOR APPLICATION NUMBER: 60/216,503
: PRIOR FILING DATE: 2000-06-30
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 125
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Consensus amino acid sequence
US-09-896-522-5

Alignment Scores:
Pred. No.: 4,97e-24 Length: 125
Score: 428.50 Matches: 84
Percent Similarity: 76.80% Conservative: 12
Best Local Similarity: 67.20% Mismatches: 28
Query Match: 14.44% Indels: 1
DB: 9 Gaps: 1

US-09-896-522-1 (1-1624) x US-09-896-522-5 (1-125)
Qy      553  CTGCGCCTCTTCGTGAGACACCGACTCCGAGCTGTCTTCGAAGAGTTCTCCGGGAC
               |||
Db      1  leuylsilepnevalaspthraspalaspvalargleuillleargrarglleuysargasp
               |||
Qy      613  GTG---CGCCGAGGAGGAGCACTTGAGCAGATTTCAGCGCAGTACACCACTTCGTGAAG
               |||
Db      21  ValaangluargglyargaspillegluservalillegluglnhrlyrleuysPheValys
               |||
Qy      670  CGCGGCTTCGAGAGAGTTCTGCTGCCGCAAGAAAGATATGCCAGTGTGATCATCCACAGA
               |||
Db      41  PrometeryrgluglnPheillegluProthrlyslstyAlaaspillellelleProarg
               |||
Qy      730  GGAAGTGACAAATATGATGCCATCAACTGATCGTGAGCAGCATCCAGGACATTTGTAAT
               |||

```

Db 61 G1G1G1AAsPAsHnH1sValAlAlleAsPLeu11leValGlnH1s1leGlnSer1leLeuAsn 80

Qy 790 GGTGATCAATCTGGCAATATGGCAGCAGGAGGCTCCAAATGGCGAGACTTACAAAGCGGACCTTT 849

Db 81 G1uG1yLeuSerSeSerg1nH1sThrAsnTyrmelValAsnArgserTyrlYsArgThPhe 100

Qy 850 TCTGAGCCAGGAGGAGCACCCTGGATGCTGACCTCTGGCAACGATTCACATTGGAGTCC 909

Db 101 SerG1uPProG1yAsPAsH1sPProG1yTyrlrThrProSerG1yLysArgGlnH1s1leuG1uSer 120

Qy 910 AGCAGCAGACCCAC 924

Db 121 SerSerArgProHis 125

RESULT 4
US-10-029-386-28765
; Sequence 28765, Application US/100293386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, Sharon R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: ABOICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28765
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALU8.00e-26
US-10-029-386-28765

Alignment Scores:
Pred. No.: 1,91e-16 Length: 60
Score: 327.00 Matches: 60
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.02% Indels: 0
DB: 12 Gaps: 0

US-09-896-522-1 (1-1624) x US-10-029-386-28765 (1-60)

Qy 745 GTTGCAATCAACCTGATGTGTGACGACATCCAGACATTTCTGAATGTGACATCTGCAA 804

Db 1 ValAlAlleAsnLeu11leValGlnH1s1leGlnAsP1leLeuAsnG1yAsP1leCysLys 20

Qy 805 TGGCAGCAGAGGAGGCTCCAAATGGCGAGACTACAAAGCGAGCCTTTCTGAGCCAGGGAG 864

Db 21 TrpHisArgG1yG1ySerAsnG1yArgserTyrlYsArgThrPheSerG1uPProG1yAsP 40

Qy 865 CACCTGGGATGCTGACCTCTGGCAACGATTCACATTGGAGTCCAGCAGCAGACCCAC 924

Db 41 HisProG1yMetLeuThrSerG1yLysArgserH1s1leuG1uSerSerSerArgProHis 60

RESULT 5
US-09-896-522-6
; Sequence 6, Application US/09896522
; Patent No. US20020055161A1

```
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 124
; TYPE: PR1
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
; US-09-896-522-6

Alignment Scores:
Pred. No.: 7.12e-16 Length: 124
Score: 320.50 Matches: 62
Percent Similarity: 67.19% Conservative: 24
Best Local Similarity: 48.44% Mismatches: 37
Query Match: 10.80% Indels: 5
DB: 9 Gaps: 2

US-09-896-522-1 (1-1624) x US-09-896-522-6 (1-124)
QY 166 CTGATAGGGGTGAGCGCGGCACTGCGAGCGGAAGTGCACCGTGTGTGAGAGATCATG 225
Db 1 llllelllellleallleaglysergylserglyshtrhrllleallleaglylleval 20
QY 226 GAGTGTCTGGACAGCAAGAGAGTGAACGCGGACGCGGAAGTGTGTCTGAGCCAG 285
Db 21 Gluwerleuasn-----lyserproglynglnlulysvallellelleseerln 36
QY 286 GACAGGTTCTACAGAGTCTCTGAGCGAGAGAGAGAGAGAGCGCTTGAAGAGACATAC 345
Db 37 Aspsantrlyrrlyrlyaspleusecgluleuaspwecgluaglysgluasnastrtyr 56
QY 346 AATTGATACCATCCAGATGCTTGTATATGATTTGATGACAGAGACTGTGAAGACATC 405
Db 57 AsnheasphlserproabpalaPheasppheaspleueutyrglnstleuys***Leu 76
QY 406 GTGAGGGCAAAACGGTGGAGGTGCCACCTATGATTTTGTGA---CACTCAAGTTA 462
Db 77 LysasnnglyluserValGluValProIleTyraSpheLysTrhLSHsArgArglys 96
QY 463 CCAAGAGCAACGGTGTCTACCCCTGCGAGCGTGTGTGTTGAGGGCATCTGTGTTTC 522
Db 97 AspeluhtrValThrlllegluProalaspallellelleuuglylletyralaleu 116
QY 523 TACAGCAGAGATCCGGGACATG 546
Db 117 TyraSpgluArglleArgaspleu 124

RESULT 6
US-09-896-522-4
; Sequence 4, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
```

```
; LENGTH: 231
; TYPE: PR1
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
; US-09-896-522-4

Alignment Scores:
Pred. No.: 1.92e-12 Length: 231
Score: 276.00 Matches: 68
Percent Similarity: 50.95% Conservative: 39
Best Local Similarity: 32.38% Mismatches: 81
Query Match: 9.30% Indels: 22
DB: 9 Gaps: 5

US-09-896-522-1 (1-1624) x US-09-896-522-4 (1-231)
QY 166 CTGATAGGGGTGAGCGCGGCACTGCGAGCGGAAGTGCACCGTGTGTGAGAGATCATG 225
Db 1 ValllelllyValalaglyserserglyalaglyllythrValalalarghrglleval 20
QY 226 GAGTGTCTGGACAGCAAGAGTGT-----GACACGGCGCAG 261
Db 21 SerlllePhegllyArggluGlyValProalalaglyllegluGlyasnProaspsersaen 40
QY 262 CGAAGGTGTCTCCTGAGCGGACGAGGTTCTCAAGGTCTCGACGGAGCAGAG 321
Db 41 ThrGlyAspsersPheleuArgleuaspArgPheTyMetaspleuhsleuGluaspArg 60
QY 322 GCCAAGGCTTGAAGAGACAGTACATTTTGTACATCCATCCAGATCCCTTGTATGATTTG 381
Db 61 LysArgAlaglyasnlyshleryserPhepaseserProgluAlasnaSpheaspleu 80
QY 382 ATCCACAGACTCTGAAGAACATCTGTAGAGGCAAAACGTTGAGGTGCCACTTGTAT 441
Db 81 LeuTyrgluValPheLysgluleuLysgluGlyLysSerValaspsProIleTyraen 100
QY 442 TTTGTACACACATCCAGATTCACGAG-----ACCAAGGTGTCTACCGCT 486
Db 101 HisValThrGlyGluLysaspProaspsrglynglnluproglyThrPheThrAsprrpro 120
QY 487 -----CGGACGTGTCTGTGTGAGGACATCTGTGTTCTACACCGCAG 531
Db 121 GluleuIllegluGlylalaaspValleuValllegluGlyleuhsalaleuTyrrasglu 140
QY 532 -----GAGTCCGGACATGTTCCACTGCGCTCTGTGACACCGATCCGAC 582
Db 141 ArgGluValasnaValalaglnleuLeuaspLeuLyslleTyraValaspsProaspleasp 160
QY 583 GTGAGGCTGTCTGAAGAGTTCGCGGACGTG---CGCGAGGAGGAGACCTGAGCAG 639
Db 161 LeuGluLeuAlaArgylslllegluArgpameAlaGluArglyhlsereleuGlyly 180
QY 640 ATTGTACGAGATPACCACTTGTGAAGCGCGCTTGAAGAGATTCGCTCGCGCAG 699
Db 181 ValLeuaspserlllegluLysArgArgLysProasPtyrValasnTyrrllealPProglu 200
QY 700 AAGAAGTATGCCATGTATCCACGA 729
Db 201 PheSerTyraLaspLeuIllelignArg 210

RESULT 7
US-10-156-761-10907
; Sequence 10907, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, UTN
; APPLICANT: HORIKAWA, HITROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATORI, MASAHIRA
```

```

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10907
; LENGTH: 797
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10907

Alignment Scores:
Pred. No.: 5,34e-06 Length: 797
Score: 192.50 Matches: 115
Percent Similarity: 35.03% Conservative: 36
Best Local Similarity: 26.68% Mismatches: 158
Query Match: 6.53% Indels: 122
DB: 15 Gaps: 22

US-09-896-522-1 (1-1624) x US-10-156-761-10907 (1-797)
Oy 1181 CTCAGTCACTAGAGGAGATCTTTAAACCGC-----AACGAGCTTAAG 1140
Db 333 IletHrbaRbLeuHleAlaAlaLeuHsnYsAlaYsAlaAbpIleMeTaSpGlyProHla 352
Oy 1139 TGGCTGAACACCTCAGAACCGCTGTCAGTGTCCCGACGAAGTGAAGTCTGAGTACAT 1080
Db 353 SerMeThIstHrAArguLeuProValIleIleGln----- 364
Oy 1079 CTGAGTTTCCACTCTGTAGTATAGGAAGGCTCGCTGCTAAACATCCCTGGGGTGGCGG 1020
Db 365 -----HIsGln-----ArgTgaIaYsThrLeuAlaLeu-ThrGlnLeuSerAl 379
Oy 1019 AGAGGAGACAGTGGGTGTGGGTGGCGCTGCCCGAGGCTCAGTCCCTGAACACATGCGCG 960
Db 379 aleuHIsAlaHrSerSerArgProArgPro-----CybG1 391
Oy 959 GCGGAGACCTGCCCTGTAGGCTCGGACGCCCTCAGTGGGGTGTGCTGTGACTCCAA 900
Db 391 YArg-----SerAlaProSerProLeuProAlaSerGly-----AlaGlnAr 406
Oy 899 TGTGACCGTTGGCCGAGAGTCAGATCCCAAGGTGTGTCCTGGGTCAGAAAGTGCAC 840
Db 406 gSerValIleCybGlyAlaSerValAlaThAlaArgHrThrGlyGlnYsGlyPheSe 426
Oy 839 TTGAGACTCGCCCATTTGACC-----CTCTCGGTGCCATTGTCAG 798
Db 426 rValSerSerThrArgValThrLeuArgArgAnsGlnArgLeuValTrpAlaLeuAlaG1 446
Oy 797 ATGTCACATTCAGAAATGCTCTGGATGTGCTGCACAGATCAGATTGATGACACATATTG 738
Db 446 nArgSerHIsValProCybAlaGlyCybPro----- 456
Oy 737 TCCACTCTCTGGGATGATCACAATCGGCATCTTTGTGTGGGAGGACAGACATCCT-- 680
Db 457 -----GluseHrArgLeuArgAlaLeuArgProLeuProSerProSe 470
Oy 679 -----CGAAGCGCGGTTTCAAGAAAGTGTGTACTGGTCAATC 639
Db 470 rLeuLeuProProAlaAlaHrArgArgProAlaSerProArg----- 483
Oy 638 TGCTCCAGATCCCTCCGTGGGAGCAGCGCCGGAAGAACTCTTGAG-----ACAGCTG 585
Db 484 ----ProHlaAlaAlaAlaAlaAlaArgProArgAlaLeuProSerProHrThrSerAl 502
Oy 584 ACGTGCAGTGCAGTGTCCACGAAGAAGGCGCAGGTGAACATGTCCTCGGATCTCTGCTG 525
Db 502 aAlaArgAlaHrSerProSerMet-----ThrLeuProProProValTrpAr 518

```

QY	524	TAGAAACACCAAGATGCCCTCAACAAGAACACAGCTCCGACGGGTAGA-----	479
Db	518	gargProAlaIArgSerSerArgProProProIArgProSerSerArgAlaIArgSerPr	538
QY	478	-----CCACCGTGGTCTTGTAACCTTGAGTGT	450
Db	538	oThrProThrArgCYSerSerValTyrSerSerSerProSerProAlaThrArgSerVa	558
QY	449	GTCACAAATCATAGGTGGGACCTTCACCGTTTGCCCTCCACGATGTTCTTCAGAGTC	390
Db	558	IserIaSer---ProThrProProProSerArgIylsPProProArgSerThr-----	574
QY	389	CTGTGCATCAATCATCTTATCAAAAGCAGCATCTGGATGTGCAAAATTGTACTGTCTTCAAG	330
Db	575	-----ArgArgProProSerAlaSerSerThrThrSerArgSerAr	568
QY	329	GCCTTGACCTTGTGCTCTGCGGTGAGACCTGTGAAACCTGTCTCGCTCAGAGTACC	270
Db	588	gArgThrThrSerArThrSerSerSerSerThrArgArgArg-----Pr	602
QY	269	ACCTTCGCGCTGCGCTGTTTCACCTCGTCTGTGTCCACGACMACATCATGATCTTTCACAC	210
Db	602	oProThrSerProAlaSerProPro-----ProArgProProSerArgThrSerSe	619
QY	209	ACGGTCGACTTCCCGCTGCGACGTGCCCGCCCTCACCCCTATAGAGAAGGGCGCTGTGTC	150
Db	619	IAlaSerSerAlaIaTPr---ThrSerArgSerThrSerSerArgArgAlaSerArgAr	638
QY	149	GGAGCGTGGCGCTCCGGCGGGGCGTCTCGCAGTGTGGCCCTCCGCGGACGACATCTGC	90
Db	638	gAlaSerArThrProAlaArgArgAlaSerArgSerSerThrPro-----Ar	655
QY	89	GCTTCGCTCCCGCGCATCGGGTCCCGCGCGCCGCCCTTCCCGGCGCCCG---CGCGC	33
Db	655	gProPro---ProArgAlaIaGlySerPro-----AlaArgThrArgAl	668
QY	32	CCGCCAGCGCCGAGGTGGAAGCGCACCCCA	2
Db	668	aArgProArgProAlaArgAlaArgSerAlaPro	678

RESULT 8

US-09-738-973-187

; Sequence 187, Application US/09738973

; Patent No. US20020110563A1

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Lodes, Michael J.

; APPLICANT: Fling, Steven P.

; APPLICANT: Mohamatch, Raodoh

; APPLICANT: Algate, Paul A.

; APPLICANT: Secrist, Heather

; APPLICANT: Indrias, Carol Yoseph

; APPLICANT: Benson, Darin R.

; APPLICANT: Elliott, Mark

; APPLICANT: Mannion, Jane

; APPLICANT: Kalos, Michael D.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR

; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.475C9

; CURRENT APPLICATION NUMBER: US/09/738.973

; CURRENT FILING DATE: 2000-12-14

; NUMBER OF SEQ ID NOS: 587

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 187

; LENGTH: 595

; TYPE: PRT

; ORGANISM: Homo sapien

US-09-738-973-187

Alignment Scores:

Pred. No.: 7.34e-05

Score: 177.00

Length: 595

Matches: 110

Percent Similarity: 33.42%
 Best Local Similarity: 27.64%
 Query Match: 5.97%
 DB: 10
 Gaps: 22

US-09-896-522-1 (1-1624) x US-09-738-973-187 (1-595)

```

QY 5 GGTGCGCTCCGACCTCGCGCTGGGCGGCGCGCGCGCGGAGAGGCGGCGCGG 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 228 GlyAlaSerAlaProGluSerGlnAlaGlyGly---GlyPro---ArgGlyArgAlaArg 245
QY 65 GGACCCGATGCGCGGAGCGGAGCGCGAGATGGCTTCGCGGAGCGGAGATGCGGAGA 124
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 GlyPro---ArgGlnGln-GlyArgArgArgHisGlyThrGlnArgAlaArgGlyPr 263
QY 125 GCCCGCGCGCGGA-----GCCCGCGCC 145
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 263 oProGlnAlaArgGluGluGlyProArgAspAlaThrThrIleuGlyLeuGlyThrPr 283
QY 146 GTCCGCA-----CCAGCG----- 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 283 oSerGlyGluGlnArgAlaAspGlnSerGlnAlaLeuProAlaLeuAlaGlyAlaAla 303
QY 159 -----GCCCTTCGATAGGGGTGAGCGCGCGCACTGCCAGCGGAAATCGA 205
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 aAlaHisAlaHisAlaIleProGlyAlaGly--ProAlaAlaAlaProValGlyGlyArg 322
QY 206 CCGTGTGTGAAGATCATGAGTGTCTGGAGCAAGAAAGGTGGAGAGCGGCGGAGA 265
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 323 Gly-----ArgArgGly-GlyTTPArgGlyGlyArgArgGlySerAlaGlyAlaG 340
QY 266 AGGTGTGATCTCTGAGCGGAGCAAGTTCTACAAGTCTCTGACGACAGAGCAAGAGCCA 325
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 340 YGlyGly-----GlyArgGlyGlyArgGlyArgGlyArgGlyArg 351
QY 326 AGGCTTGAAGAGCAGTACATTTTGACCATCCAGATCCTTGATATGATTTGATGC 385
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 351 gGly----- 352
QY 386 ACAGAGACTGTGAAGAACATCGTGGAGGCAAAAGGTGGAGTGCAGCCTATGATTTTG 445
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 353 -----GlyGlyArgGlyGlyGlyAla 360
QY 446 TGACACTCAAGTTACAGAGACACAGTGTCTACCTGCGAGCGTGTCTGTTTG 505
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 -----GlyArgGlyGly----- 364
QY 506 AGGCGATTTGTTGTTCTACAGCCAGAGATCCGGAGATGTTCCACTGCGCTCTTCG 565
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 365 -----GlyAlaAlaGlyProArgGly---GlyAlaSerSerProGlyAlaArgAr 380
QY 566 TGAACACCGACTCCGACGTCAGGCTGTCTGAAAGATTCCTCGGAGAGTGGCGGAGAGA 625
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 380 gGlyGluGlnArgArgArg-----GlyArgGlyPro---Pr 391
QY 626 GGAAGCTGGAGCAGATTCTGACGAGTACACACACTTCGTGAAAGCGGCTTCGA----- 680
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 391 oAlaAlaGlyAlaAlaGlnValSerAlaArgGlyArgArgAlaArgGlyGlnAlaG 411
QY 681 -----GGAGTTTGCCTGCGGACCAAGAGATAGCCGATGTGATTC 724
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 411 yGluGlnAlaGlnAspGlyLeuLeuProArgGlyArgAspArgLeu-----Pr 427
QY 725 CACGAGAGTGTGACAAATGTGGTGCATCACTGATGTCACACACATCCAGGACATTC 784
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 427 oLeuArgProGlyAspAlaAsn-----GlnArgAlaGluArgProGly----- 441
QY 785 TGAATGTGACATCTGCAATGTGACACCGAGAGGAGTCCATAGG-----CGAGCTACA 838
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 -----ProProArg-GlyGlyHisGlyProValAspAlaSer 454
QY 839 AGCGGACCTTTTCTGAGCCAGGAGACCACTTGGAGATGCTGACTCTGGCAACGCTCAC 898
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 454 eAlaProAspThrSerProProAlaHisProAlaArgTrpValSerGlnAlaArgGlnA 474
QY 899 ATTTGAGTGTCCAGACAGACCCCACTGAGAG--GCTGCCAGCGCTCAGGAGGTCTC 955
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 474 rGluTTPArgGlnPheArgValGlyGlyGlyPheProProProProSerArgProP 494
QY 956 CCGCC-----CGCATGTGTTCAGGAGTACGACCTTGGAGCGCC 997
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 494 roAlaValLeuLeuProLeuLeuArgLeuAlaCysAlaGly---AspProGlyAlaThrA 513
QY 998 ACCCACACCACTGCTTCCTTCGCGCGACCCCAAGGAGAGTGT 1041
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 513 rGProGlyProArgArgProAlaArgArgProArgGlyGluLeu 527

```

RESULT 9

```

US-09-854-133-187
; Sequence 187, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 187
; LENGTH: 595
; TYPE: PR
; ORGANISM: Homo sapien
US-09-854-133-187

```

Alignment Scores:

```

Pred. No.: 7,34e-05 Length: 595
Score: 177.00 Matches: 110
Percent Similarity: 33.42% Conservative: 23
Best Local Similarity: 27.64% Mismatches: 118
Query Match: 5.97% Indels: 147
DB: 10 Gaps: 22

```

US-09-896-522-1 (1-1624) x US-09-854-133-187 (1-595)

```

QY 5 GGTGCGCTCCGACCTCGCGCTGGGCGGCGCGCGGAGAGGCGGCGCGG 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 228 GlyAlaSerAlaProGluSerGlnAlaGlyGly---GlyPro---ArgGlyArgAlaArg 245
QY 65 GGAAGCTGGAGCAGATTCTGACGAGTACACACACTTCGTGAAAGCGGCTTCGA----- 124
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 GlyPro---ArgGlnGln-GlyArgArgArgHisGlyThrGlnArgArgGlyPr 263
QY 125 GCCCGCGCGCGGA-----GCCCGCGCC 145
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 263 oProGlnAlaArgGluGluGlyProArgAspAlaThrThrIleuGlyLeuGlyThrPr 283
QY 146 GTCCGCA-----CCAGCG----- 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 283 oSerGlyGluGlnArgAlaAspGlnSerGlnAlaLeuProAlaLeuAlaGlyAlaAla 303
QY 159 -----GCCCTTCGATAGGGGTGAGCGCGCGCACTGCCAGCGGAAATCGA 205
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 aAlaHisAlaHisAlaIleProGlyAlaGly--ProAlaAlaAlaProValGlyGlyArg 322
QY 206 CCGTGTGTGAAGATCATGAGTGTCTGGAGCAAGAAAGGTGGAGAGCGGCGGAGA 265
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 323 Gly-----ArgArgGly-GlyTTPArgGlyGlyArgArgGlySerAlaGlyAlaG 340
QY 266 AGGTGTGATCTCTGAGCCAGGAGACAGGTTCTACAAGTCTCTGACGACAGAGCAAGAGCCA 325
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db      340 yGlyGly-----GlyArgGlyGlyArgGlyArgGlyArg 351
Qy      326 AGGCCTTGAAGAAGACATGATTTTGACATCCAGATGCTTTGATTAATGATTTGATGC 385
Db      351 gGly----- 352
Qy      386 ACAGAGCTCTGAAGAATCGTGAGGGCAAAACGGTGAGTGCCGACCTATGATTTTG 445
Db      353 -----GlyGlyArgGlyGlyGlyGlyGlyA----- 360
Qy      446 TGACACACTCAAGTTACAGAGACCAAGGTGTTACCTCGCGAGCGTGTCTGTTTG 505
Db      361 -----GlyArgGlyGly----- 364
Qy      506 AGGCATCTTGTTGTTCTACAGCAGAGATCCGGACATGTTCCACTTCGCTCTTCG 565
Db      365 -----GlyAlaAlaGlyProArgGlu--GlyAlaSerSerProGlyAlaArgAr 380
Qy      566 TGACACCGACTCCGAGCTGCTCGAAGAGTTTCGGGAGCTGCGCGAGGA 625
Db      380 gGlyGlyGluArgArgArg-----GlyArgGlyPro--Pr 391
Qy      626 GGGACCTGAGCAGATTCTGACGACGTACACACTTCGTGAAGCCGCGCTTGA----- 680
Db      391 oAlaAlaGlyAlaAlaGlnValSerAlaArgGlyArgArgAlaArgGlyGlnArgAlaG 411
Qy      681 -----GGAATTCTGCTGCGCGCAAAAGATAGCCGATGTGATCATCC 724
Db      411 yGlnGluAlaGlnAspGlyLeuLeuProArgGlyArgAspArgLeu-----Pr 427
Qy      725 CACGAGAGTGACCAATATGTTGCCATCAACCTGATCGTGACGACATCCAGGACATTC 784
Db      427 oLeuArgProGlyAlaAlaAsn-----GlnArgAlaGlnArgProGly----- 441
Qy      785 TGAATGTGACATCTGCAAAATGACACCGAGAGGTCATGCG-----CGAGCTACA 838
Db      442 -----ProProArg-GlyGlyHisGlyProValAlaAlaSerS 454
Qy      839 AGCGGACCTTTTTCAGCCAGGGAGCAACCTTGAGTCTGACCTCTGGCAAAACGGTCA 898
Db      454 eAlaProAspThrSerProArgHisAspProArgTrpValSerGlnGlnArgGlnA 474
Qy      899 ATTGGAGTCCAGACAGACCCCACTGAGAG--GCTGCCAGAGCTTCAGGAGCGCTC 955
Db      474 rGluTrpArgGlnPheArgValGlyGlyPheProProProProSerArgProp 494
Qy      956 CCGCC-----CGCATGTGTTCAGGAGTGAAGCTTGAGGAGCGCC 997
Db      494 rAlaValLeuLeuProLeuLeuArgLeuAlaCysAlaGly--AspProGlyAlaThrA 513
Qy      998 ACCCAGACCACTGCTTCTCTCGGCGCACCCCAAGGGGAGTGT 1041
Db      513 rGProGlyProArgArgProAlaArgArgProArgGlyGluLeu 527

RESULT 10
US-10-144-649A-187
; Sequence 187, Application US/10144649A
; Publication No. US2003011859A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Pan, Liqun
; APPLICANT: Algate, Paul A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C11
; CURRENT APPLICATION NUMBER: US/10/144,649A
; NUMBER OF SEQ ID NOS: 749
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 187
; LENGTH: 595

```

```

; TYPE: PRT
; ORGANISM: Homo sapien
US-10-144-649A-187

Alignment Scores:
Pred. No.: 7,34e-05 Length: 595
Score: 177.00 Matches: 110
Percent Similarity: 33.42% Conservative: 23
Best Local Similarity: 27.64% Mismatches: 118
Query Match: 5.97% Indels: 147
DB: 15 Gaps: 22

US-09-896-522-1 (1-1624) x US-10-144-649A-187 (1-595)

Qy      5 GGTGCGCTCCGACCTGCGGCTGGGCGCGCGCGCGCGGAGAGGCGCGCGCG 64
Db      228 GAlaAlaSerAlaProGlnSerGlnAlaGlyGly--GlyPro--ArgGlyArgAlaArg 245
Qy      65 GGAACCGGATGCGCGGAGCGGAGCGCGGAGATGAGCTTCGCGGAGGCGAAGACTCGGAGA 124
Db      246 GlyPro-----ArgGlnGln-GlyArgArgArgHisGlyThrGlnArgArgGlyPr 263
Qy      125 GCGCCGCGCGCGA-----GCGCGACC 145
Db      263 oProGlnAlaArgGlnGluGluGlyProArgAspAlaThrThrIleLeuGlyLeuGlyThrPr 283
Qy      146 GTCCGCA-----CGAGG----- 158
Db      283 oSerGlyGlnGlnArgAlaAspGlnSerGlnAlaLeuProAlaLeuAlaGlyAlaAla 303
Qy      159 -----GCCCTTCGTATGAGGATGAGCGCGCGCACTGCGAGCGGAGATCGA 205
Db      303 aAlaHisAlaHisAlaIleProGlyAlaGly--ProAlaAlaAlaProValGlyGlyArg 322
Qy      206 CCGTGtGTAAGAATCATGAGTCTGCGGACAGAACAGAGTGAACAGCGCGACCGGA 265
Db      323 Gly-----ArgArgGly-GlyTrpArgGlyArgArgGlyGlySerAlaGlyAlaG 340
Qy      266 AGGTGTCTACTCTGAGCCAGGACAGGTTCTACAAGTCTGACGCGACAGAGAGAGCA 325
Db      340 yGlyGly-----GlyArgGlyGlyArgGlyArgGlyArg 351
Qy      326 AGGCCTTGAAGAAGACATGATTTTGACATCCAGATGCTTTGATTAATGATTTGATGC 385
Db      351 gGly----- 352
Qy      386 ACAGAGCTCTGAAGAATCGTGAGGGCAAAACGGTGAGTGCCGACCTATGATTTTG 445
Db      353 -----GlyAlaAlaGlyProArgGlu--GlyAlaSerSerProGlyAlaArgAr 380
Qy      446 TGACACACTCAAGTTACAGAGACCAAGGTGTTACCTCGCGAGCGTGTCTGTTTG 505
Db      361 -----GlyArgGlyGly----- 364
Qy      506 AGGCATCTTGTTGTTCTACAGCAGAGATCCGGACATGTTCCACTTCGCTCTTCG 565
Db      365 -----GlyAlaAlaGlyProArgGlu--GlyAlaSerSerProGlyAlaArgAr 380
Qy      566 TGACACCGACTCCGAGCTGCTCGAAGAGTTTCGGGAGCTGCGCGAGGA 625
Db      380 gGlyGlyGluArgArgArg-----GlyArgGlyPro--Pr 391
Qy      626 GGGACCTGAGCAGATTCTGACGACGTACACACTTCGTGAAGCCGCGCTTGA----- 680
Db      391 oAlaAlaGlyAlaAlaGlnValSerAlaArgGlyArgArgAlaArgGlyGlnArgAlaG 411
Qy      681 -----GGAATTCTGCTGCGCGCAAAAGATAGCCGATGTGATCATCC 724
Db      411 yGlnGluAlaGlnAspGlyLeuLeuProArgGlyArgAspArgLeu-----Pr 427
Qy      725 CACGAGAGTGACCAATATGTTGCCATCAACCTGATCGTGACGACATTC 784
Db      427 oLeuArgProGlyAlaAlaAsn-----GlnArgAlaGlnArgProGly----- 441

```


[illegible][illegible]

```

Db 143 -----ProArgTyrGlyProLeu-----1le 149
QY 887 CCAGAGTCAGCATCCAGGGTGGTCCCTGAGTCCAGAAAGTCCGCTTAGCTCCGC 828
Db 150 ProGluHisSer-----SerGlyHisProLysSerSerProValAla-----PheArg 165
QY 827 CCATTGGACCTCTCT-----CGGTGCCATTGGAGATGTCACCATTCAGAAATG 780
Db 166 ProLeuHisCysProPheLeuGluThrLysIleLeuGluArgAlaProPhe-----183
QY 779 TCCGCGATG-----TGCTGCACGATCAGGTGATGCGAAC-----744
Db 184 ---ThrValProThrCysLeuProProTyrIleMetSerSerLeuProProGluArgSer 202
QY 743 -----ATTATGTCCACTCCCTCGGGATGATGATCA 714
Db 203 TyrAspTyrProLeuAlaProSerProTyrValTyrSerGlySerGlnProLysValPro 222
QY 713 TCGGCATACCTTTTGTGCGCA-----GGCAGAACTCTCGA-----AGGCC 672
Db 223 SerAlaPhe-SerLeuGlySerLysGlyPheTyrHisLysAspProAsnIleLeuArgPr 242
QY 671 GGCCTTCAGAAAGGTGTACTGGCTAGAAATCT-----638
Db 242 oAla---LysGluProLeuAlaAlaSerGlySerGlyMetLeuGlyLeuAlaProGlyG1 261
QY 637 -----GCTCCAGATCCCTCCCTCGGCGACAGTCCCGAGAACT 600
Db 261 yHisLeuGlnGlnAlaCysAspAlaGluGlyProSerIleuHisGlnArgAspGlyGlu-- 280
QY 599 CTTGGAAGACGCTGACGTCGGAATCGGTTCAGAAAGCGCAGGT-----GG 549
Db 281 -----ThrGlyAlaGlyArgGlnAs 288
QY 548 AACATGCCCGGATCTCTGCTGTAGAACACCAAGATCCCTCAACAGAACCAAGCTCC 489
Db 288 nLeuCysProValPheLeuGly-----TyrProAspThrValProArgGln 303
QY 488 G-----CAGGGTAGACCAACCGTGTCTGTGTAACCTTG-----455
Db 303 rProTyrProSerLysProProGlyGlyCysSerSerHis-----410
QY 454 ---AGTGTTCACAAATCATAGTCGCGACCTCA-----CCGTTTGCCTT-----410
Db 323 oGlySerAsnSerPheGlyTyrGlnLeuGlyProProValThrProArgCysProSerPr 343
QY 409 -----CSACAGATGTTCTTCAAGAGTCCCTGTGCATCAAACTATTAAGGATCTGATGG 354
Db 343 oGlyProProThrProProGlyGlyCysSerSerHis-----356
QY 353 TCAAAATTGTAAGTCTCTTCAAGGCTTGGCTTGTCTGCTGCGCTCA-----305
Db 357 -----LeuProAlaArgGluGlyAs 363
QY 304 ---GAACTTTGTAACCTGTCTGCTGCTAGATGACACCTTCCGCTGCGCTGTTCC 249
Db 363 rProGlyProCysArgLysCysGlnAspSer-----ProGluGlySerSerSerGlyPr 381
QY 248 ACCTGCTGTGTGCCA-----GCAACTCCAGATCTTCTCA 213
Db 381 oGlyGluSerSerGluGluThrArgAsnLysAlaGlySerArgAlaSerProProSerHisH1 401
QY 212 CACACGCTGACATTCGCGCTGAGTGCAGTCCGCGCTCAACCTTATCAGAGAGGCGCGCTGG 153
Db 401 sThrLysLeuLysLysThrLysThrLysThrLysHisSerGluGlnPheGluCysProGlyG1 421
QY 152 TCGGAGAGCTGCG-----CCTCGGCGCGGAGCTTCGAGATCTTCCGCTCCGCGGAA 99
Db 421 yCysProGlyLysGlyLysSerProAlaThrGlyLeuArgAlaLeuLysArgAlaGlySe 441
QY 98 GCGA-----TCTGGGCTCGGCTCCGCGGATCGGGTCCCGCGCGCGCGCT 51
Db 441 rProGluValGlnGlyValaArgGlyProAlaProLys-----ArgPro-SerHisThrP 459

```

```

QY 50 TCCCGGAGCCGCGCGCGCCGCGGAGCGGAGTGGAGGCGACCC 4
Db 459 heProGlyThrGlyArgGlnGlyAlaArgAlaTrpGlnGlnThrPro 474

RESULT 14
US-10-078-547-24
; Sequence 24, Application US/10078547
; Publication No. US20020199211A1
; GENERAL INFORMATION:
; APPLICANT: Narayanaswamy Ramesh
; APPLICANT: Miguel A. de la Fuente
; APPLICANT: Ines M. Anton
; APPLICANT: Ralf S. Geha
; TITLE OF INVENTION: WIP. A WASP-Associated Protein
; FILE REFERENCE: 1242.1022-005
; CURRENT APPLICATION NUMBER: US/10/078,547
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 09/599,287
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: PCT/US98/27501
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/101,457
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/068,533
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ. ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ. ID NO 24
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translated WIP ORF No. US20020199211A1 3
US-10-078-547-24

Alignment Scores:
Pred. No.: 0.00676 Length: 507
Score: 164.00 Matches: 99
Percent Similarity: 34.41% Conservative: 18
Best Local Similarity: 29.12% Mismatches: 156
Query Match: 5.57% Indels: 67
DB: 14 Gaps: 15

US-09-896-522-1 (1-1624) x US-10-078-547-24 (1-507)
QY 1036 TCCCTGGGGTGGGCGGAGAGAGACATGGGTGGTGGGCTCCCAAGCTCACTCC 977
Db 157 SerProGlyHisArg-----SerGlyProProGluProGlnArg 169
QY 976 CTGACACACATCCGCGGCGGAGACCTGCTGAGAGCTCGGACGCCCTCAGTGGGGTC 917
Db 170 ---Asn-ArgMetProProProArgProAspValGlySerLysProAspSerL1eProPr 188
QY 916 TGCTGTGATCTCAAAATGTGACCGTTTGCACAGAGTCAAGATCCGAGGTGTCCCTG 857
Db 188 oProValProSerThrProArgProL1eGlnSerSerL1eHisAsnArgGlySerProPr 208
QY 856 GCTCAGAAAAAGTCCGCTGTAGCTCCGCCATTTGACCTCTCGTGGTCCATTGGCAGA 797
Db 208 oValProGlyGlyProArgGlnProSerPro--GlyProThrProProPhe-----225
QY 796 TGTACACATTCAGAATGTCTGATGTGTGACATCAGTTGATGGCAACCATATTGT 737
Db 226 -----ProGlyAsnArg--Gly 230
QY 736 CCACTCTCTGTGAGATGATCACATCGCACTTCTTTGTG-----GCA 692
Db 231 ThrAlaLeuGlyGlyCylSerL1eArgGlnSerProLeuSerSerSerSerProPheSer 250
QY 691 GCGACAACTCTTCAAGAGCGCGCTTCACGAAGGTGTGTACTGGCTCAGAAATCTGCTCA 632
Db 251 AsnArgProProLeuProProThrProSerArgAlaLeuAspAspLysProProProPro 270

```

```

QY 631 GGNCCCTCCCTGGCGGACGCTCCGGAGAACTTTCGAGACAGCCGTCGAGTCCG 572
DB 271 ProProProVal1GlyAsnArgPro-----Ser1LeHisArgGluAla 284
QY 571 TGTCCAGAGAGAGCCAGGTGGACATGTCGCCGATCTCTGGCTGTAGAACACAGA 512
DB 285 ValProProProProProGlnManAsnLysProProValProSer-----ThrProArg 302
QY 511 TGCCTTCMAACAGAACCGCTCCGACGGGTAGAACACCGCTGTCTGTGTAACCTTGAGT 452
DB 303 ---ProSerAlaProHisArgProHisLeuArgProPro----- 314
QY 451 GTCTCAAAAATATATATGTCGCGACCTCCACCGTTTTCGCTCAGACATTTCTTCAGAG 392
DB 315 -----ProProSerArgProGlyArgProProLeu---ProProSerSerArgLysAsn 331
QY 391 TCCTGTGCATCAATATATATCAAGGACATGTGATGTCMAAATTGTACTGTCTTTCAC 332
DB 332 AspGluThrProArgLeuProGlnArgAsnLeuSerLeuSerSerThr----- 348
QY 331 AGGCTTGGCTTCTGCTGTGCTGCGTACAGACTTGT---AGAACTGTGCTGTGCTACAGA 275
DB 349 ProProLeuProSerProGlyArgSerGlyProLeuProProProValProSerGluArg 368
QY 274 TGAACACCTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 215
DB 369 ProProProProValArgAsnProProGlyArgSerGlyProLeuProProProProPro 388
QY 214 CACACAGGTCGACTCCCGCTGGACGTGCGCGCTCACCCCTATCAGAGAGGCGGCT 155
DB 389 ValSerArgAsnLysLeu-----ThrSerArgLysLeu 399
QY 154 GGTGGAGAGCGTGGCGCTCGGCGCGG---GGCTCTTCGAGTTCGCTTCGCGCGGAG 98
DB 400 ProAlaThrProGlnLeuProSerArgSerGlyValAsnSerProArgSerGlyProArg 419
QY 97 CCATCTCGGCTCGG---CTCCCGCGCATGGGTCCCGCGCGCGCGCGCTCC 47
DB 420 ProProLeuProProAspArgProSerAlaGlyAlaProProProProPro 437

RESULT 15
US-10-017-161-1350
; Sequence 1350, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; PRIOR FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1350
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (2)..(35)
; OTHER INFORMATION: Variable amino acid
US-10-017-161-1350

Alignment Scores:
Pred. No.: 0.000744 Length: 737
Score: 164.00 Matches: 148
Percent Similarity: 33.16% Conservative: 44
Best Local Similarity: 25.56% Mismatches: 209

```

```

Query Match: 5.57% Indels: 178
DB: 12 Gaps: 29
US-09-896-522-1 (1-1624) x US-10-017-161-1350 (1-737)
QY 1339 ACCAGATCAGATGTGAAAAAATCTCTCCACTGTGGGTTCACCTGCAAAAA-----A 1286
DB 172 ThrProGlnSerThrLysProLeuSerProGlyAsnProSer--ValArgLysProLeuSer 191
QY 1285 CATCAGGCGCAGCGAGTGTCT-----AGCGTGTCTCTCAATTTCCCAATAATGTG 1235
DB 191 erProGlySerProLeuSerProArgLysProLeuSerProGlyAsnProSerValProG 211
QY 1234 CCTCAATCTCCACAGAACGCT-----CCGAGGCTCTCCGACATTTG 1190
DB 211 lYThrProGlnSerGlyAsnProSerValProGlyAsnProSerValProGlyAsn---- 229
QY 1189 TGGCATTTTCTCAGTGCCTAGAGGATCTTTAAACGCAACGACGCTAAGTGGTGAAAA 1130
DB 230 -----ProSerValProGlyThrProGlnSerArgLysPro-----ProG 243
QY 1129 CCTCAGAAAGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1076
DB 243 lnsSerGlyAsnPro--SerValProGlyThrProGlnSerGln--GluThrProGlnSerAr 262
QY 1075 GTTTCACCTCTCAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1027
DB 262 GlySerProLeuSerProArgAsnProPheSerProArgAsnProLeuGlyProGlyAsnDr 282
QY 1026 -----TGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 989
DB 282 oleuSerGlnGluThrProGlnSerArgGluProProGlnSerGlyAsnProSerValPr 302
QY 988 CAGG----- 985
DB 302 olArgAsnProLeuSerProGlyAsnProSerAlaProGlyThrProGlnSerLeuGluPr 322
QY 984 -----CTGAGTCCCTGAACACACACAGCCGCGGAGAGACCTGCCCT 944
DB 322 oleuSerProGlyLysProLeuSerProGlyAsnProPheAsnProGlyAsnProSerVa 342
QY 943 GAGGCTC----- 937
DB 342 lArgLysProLeuSerProGlyAsnProLeuSerProGlyLysProLeuSerProArgAs 362
QY 936 -----GGCAGCCCTCAAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899
DB 362 nProSerValProGlyAsnProSerValProGlyThrPro--GlnSerArgLysProPheS 382
QY 898 GTGACCGTTTGCAGAGGTCAGACATCCAGGGGTGTCCTGCTGCTGCTGCTGCTGCTGCTGCT 839
DB 382 erProArgAsnProSerValSerGlyThrProSerValProGlyThrProGlnSerArgL 402
QY 838 TGTAGCTCCGCCATTTGACCT-----CCTC 812
DB 402 yeProLeuSerProGlyAsnProPheGlyProGlyAsnProLeuValProGlyAsnProS 422
QY 811 GGTGCCATTTGCAGATGTCACCATTCAGAAATGCTCGATGTGTCGACGATCAGGTTGA 752
DB 422 erValArgLysProLeuSerProGlyAsnPro-----L 433
QY 751 TGGCAACCATATTTGTCATCTCTGTGGATATATACATCGGACATCTTTGTGGCA 692
DB 433 euser--ProGlyLysProLeuSerProGlyAsnProSer--ValArgLysProLeuSerPro 452
QY 691 GGCAGA-----ACTCCTGGAAGCGCGCTTGCAGAGAGGTGTACTGCTGACAGATCT 638
DB 453 GlyLysProLeuSerProArgLysProLeuSerProGlyAsnProGlnSerArgLysPro 472
QY 637 ---GCTCAGGT-----CCCTTCCTGGGCGCAGTCCCGGAGACTTTTCGAG 593
DB 473 LeuSerProGlyLysProLeuSerProGlyLysProLeuSerProGlyAsn----- 489

```

```
QY 592 ACAGCTGACGTCGAGTCGCTCCACGAAGAGCGCAGGTGGAACATGTCCCGATCT 533
Db 490 -----ProleuserProGlyLysProLeuCysProGlyAsn 501
QY 532 CCTGCTGTAGACACCAAGATGCCCT-----CAAACGAACCAAGTCGCG 488
Db 502 ProSerValArgLysProLeuserProGlyGluThrProGlnSerArgGluProLeuser 521
QY 487 CAGGGTAGACCA-----CCGTGCTCTGTACCTTGACCTGAGTGTGACAA 443
Db 522 ProGlyLysProLeuserProGlyAsnProGlnSerArgGluProLeuserProGlyLys 541
QY 442 -----AATCATAGTCGACACTCCACCGTTT 416
Db 542 ProLeuserProGlyLysProLeuserProGlyAsnProLeuserProGlyLysProLeu 561
QY 415 TGCCCTCCAGCATGTTCTTCAGAGTCCTGTCATCAATCATTCATAAGGCACTTGAT 356
Db 562 CysProGlyAsnProSerValArgLysProLeuserPro-----GlyLysProLeuser 579
QY 355 GGTCAAAATTGTAAGTCCTTCAAGGCTTGACCTTGCTGCTGCGTCAGACCTTGT 296
Db 580 ProGlyAsnProSerValProGlyAsnProSerValProGlyThrProSerValProGly 599
QY 295 AGA---ACCTGTCCTGGCTCAGATGACCA-----CCTTCCGCT 260
Db 600 AsnProSerValProGlyThr-----ProGlnSerArgAsnProSerValProGlyAsn 617
QY 259 GCCGCTGTTCCACCTCGT---TCTGTCCAGCACTCCATGATCTTCTCACACAGGTCG 203
Db 618 ProSerValProGlyAsnProSerValProGlyThrPro-----GlnSer 632
QY 202 ACTTCCCGCTGGCAGTGCAGCGCGCTCACCCCTATCAGGAAGCGCGCTGTGCGACGCT 143
Db 633 GlnGluProProGlnSerArgGluProProGlnSerArgGluThrProAlaValProGly 652
QY 142 CGGCTCCGGGCGGGGCTCTCGCAGTCTTGCTCCGCGCGAAGCAATCTCGGCTCCG 83
Db 653 AsnPro-----SerValArgLysProLeuserProGlyLysProLeu 666
QY 82 CTCCCGCGC-----ATCGGGTCCCGCGCGCCGCCCTT---CCCGGGCGCGG 38
Db 667 SerProArgLysProLeuserProGlyThrProSerValProValThrProGlnSerArg 686
QY 37 -----CGCGCCCGCCGACGCGCAGGTGCGAGGCGACCCCA 2
Db 687 GluThrProGlnSerArgGluProProGlnSerArgGluThrPro 701
```

Search completed: November 25, 2003, 07:50:17
Job time : 187.943 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2003, 07:27:32 ; Search time 20.812 Seconds

(Without alignments)
6603.180 Million cell updates/sec.

Title: US-09-896-522-1

Perfect score: 2967

Sequence: 1 gfgggggtcgccgacccgaccc.....ccaggtctgtgggggacag 1624

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues
Total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=x1h
-Q=/cgn2_1/USPTO.spool/US09896522/runat_21112003.184106.2885/app_query.fasta_1.2830
-DB=Issued_Patents_AA -QPMT=fastaan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTMT=ptco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09896522.@CGN_1.1.29 @runat_21112003.184106.2885 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGIOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_AA.*
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1363	45.9	260	US-09-536-647-2	Sequence 2, Appl1
2	1325	44.7	277	US-09-536-647-3	Sequence 3, Appl1
3	451	15.2	216	US-09-134-001C-3618	Sequence 3618, Ap
4	434	14.6	210	US-09-107-532A-5024	Sequence 5024, Ap
5	412	13.9	222	US-09-198-452A-783	Sequence 783, App
6	204.5	6.9	663	US-09-252-991A-130843	Sequence 30843, A
7	203	6.9	369	US-09-252-991A-25394	Sequence 25394, A
8	196	6.7	568	US-09-252-991A-23264	Sequence 23264, A
9	194	6.5	664	US-09-252-991A-31116	Sequence 31116, A
10	188.5	6.4	341	US-09-252-991A-32424	Sequence 32424, A
11	188	6.3	375	US-09-252-991A-31128	Sequence 31128, A
12	186.5	6.3	1427	US-09-252-991A-20577	Sequence 20577, A

C 13	186	6.3	300	4	US-09-252-991A-27679	Sequence 27679, A
C 14	185	6.3	258	4	US-09-252-991A-22452	Sequence 22452, A
C 15	184.5	6.3	1706	4	US-09-252-991A-31760	Sequence 31760, A
C 16	182	6.1	407	4	US-09-252-991A-32423	Sequence 32423, A
C 17	181.5	6.2	319	4	US-09-252-991A-32635	Sequence 32635, A
C 18	181.5	6.2	335	4	US-09-252-991A-23674	Sequence 23674, A
C 19	181.5	6.2	335	4	US-09-252-991A-24899	Sequence 24899, A
C 20	178.5	6.1	395	4	US-09-252-991A-22433	Sequence 22433, A
C 21	178.5	6.0	1706	4	US-09-252-991A-31760	Sequence 31760, A
C 22	177.5	6.0	282	4	US-09-252-991A-26169	Sequence 26169, A
C 23	177.5	6.0	313	4	US-09-252-991A-27480	Sequence 27480, A
C 24	177.5	6.0	425	4	US-09-252-991A-19214	Sequence 19214, A
C 25	177.5	6.0	467	4	US-09-252-991A-18296	Sequence 18296, A
C 26	177	6.0	266	4	US-09-252-991A-19128	Sequence 19128, A
C 27	177	6.0	437	4	US-09-252-991A-23739	Sequence 23739, A
C 28	177	6.0	595	4	US-09-370-838-187	Sequence 187, App
C 29	176.5	6.0	398	4	US-09-252-991A-26217	Sequence 26217, A
C 30	176.5	6.0	511	4	US-09-252-991A-26078	Sequence 26078, A
C 31	176.5	5.9	1113	4	US-09-252-991A-29215	Sequence 29215, A
C 32	176	6.0	472	4	US-09-252-991A-31978	Sequence 31978, A
C 33	175	5.9	582	4	US-09-252-991A-24055	Sequence 24055, A
C 34	174.5	5.9	274	4	US-09-252-991A-25569	Sequence 25569, A
C 35	174.5	5.9	631	4	US-09-252-991A-30909	Sequence 30909, A
C 36	174	5.9	904	4	US-09-252-991A-25286	Sequence 25286, A
C 37	174	5.9	977	4	US-09-252-991A-16655	Sequence 16655, A
C 38	173	5.8	590	4	US-09-252-991A-19127	Sequence 19127, A
C 39	172.5	5.9	250	4	US-09-252-991A-27781	Sequence 27781, A
C 40	172.5	5.9	286	4	US-09-252-991A-30343	Sequence 30343, A
C 41	172	5.8	316	4	US-09-252-991A-25345	Sequence 25345, A
C 42	172	5.8	333	4	US-09-252-991A-19956	Sequence 19956, A
C 43	172	5.8	391	4	US-09-252-991A-27211	Sequence 27211, A
C 44	172	5.8	683	4	US-09-252-991A-29859	Sequence 29859, A
C 45	172	5.8	2294	4	US-09-252-991A-17231	Sequence 17231, A

ALIGNMENTS

RESULT 1
US-09-536-647-2
; Sequence 2, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
; TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FaastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 260
; TYPE: PRT
; ORGANSIM: Human
US-09-536-647-2

Alignment Scores:

Pred. No.: 1,83e-112
Score: 1363.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 45.94%
Length: 260
Matches: 260
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-09-896-522-1 (1-1624) x US-09-536-647-2 (1-260)

QY 94 ATGGCTTCGAGCGGAGGCGAGACATCGGACCCCGGAGGCGGACCGTCGCGAC 153
DB 1 MetAlaSerAlaGlyGlyGluAspCysGlnSerProAlaProGluAlaSprProHis 20
QY 154 CAGCGGCGCTTCCTGATAGAGGTGAGCGGCGGACCTGCGAGGAGAGTGCACCGTGT 213
DB 21 GlnArgProPheLeuIleIleValSerGlyGlyThrAlaSerClyLysSerThrValCys 40

```

QY      214 GAGAGATCATGAGTTGCTGGGACAGAACGAGGTGGAACGCGGACCGGAAAGTGTGTC 273
Db      41 GlnuyllelmetGlnleuLeuGlnlnAsnGlnValGlnGlnArglnArglnValVal 60
QY      274 ATCTGAGCCAGGACGAGTTCTACAGAGTCTTGACCGGACAGGAAGGACGAGCTTG 333
Db      61 lLeuSerGlnAspArgPheTyrLysValAlaLeuThrAlaGlnGlnLysAlaLysAlaLeu 80
QY      334 AAAGACAGTACATATTTGACCATCCAGATGCTTTGATATATGATTTGATGACAGGACT 393
Db      81 LyselGlnTyrAsnPheAspPheProAspAlaPheAspAsnAspLeuMetHisArgThr 100
QY      394 CTGAAGACATCTGTGGAGGGGAAAACGGTGGAGGTGCCGACCTTATGATTTTGTGACAC 453
Db      101 LeuLysAsnIleValGlnGlnLysThrValGlnValProThrTyrAspPheValThrHis 120
QY      454 TCAAGTTTACAGAGACACGAGTGTGTCACCTGCGGACGAGTGTCTGTTTGAAGGACATC 513
Db      121 SerArgLeuProGlnThrThrValValTyrProAlaAspValAlaLeuPheGlnGlyIle 140
QY      514 TTGGTGTTCACAGACGAGAGATCCGGACATGTTCCACCTGCGCTCTTGTGACACCC 573
Db      141 LeuValPheTyrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
QY      574 GACTCCGACCTGAGGCTGTCTCGAAGAGTTCTCCGGACGTCGCCCGAGGAGGACCTG 633
Db      161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
QY      634 GAGAGATTCTGACGACGATACACACCTTCGTGAAGCGGCGCTTCAGAGAGTTGCTGCTG 693
Db      181 GlnGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGlnGlnPheCysLeu 200
QY      694 CCGACAAAGAAGTATGCCGATGTGATCATCCACGAGAGTGGACAAATATGTTGCCATC 753
Db      201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
QY      754 AACCTGATCTGCGACGACATCCAGACATTTGAAATGTGACATCTGCAATATGGACCGA 813
Db      221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlnLysAspIleCysLysThrPheArg 240
QY      814 GAGAGGTCCAAATGGGCGGAGCTTCAAGCGGACCTTTCTGAGCGAGGAGCAACCTGGG 873
Db      241 GlnGlnSerAsnGlnLysSerTyrLysArgThrPheSerGlnProGlnLysAspHisProGly 260

```

RESULT 2
US-09-536-647-3

; Sequence 3, Application US/09536647
; Patent No. 6579708

; GENERAL INFORMATION:

; APPLICANT: Ho, Yen Sen

; APPLICANT: Johnson, Randall

; TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase

; FILE REFERENCE: GP50020

; CURRENT APPLICATION NUMBER: US/09/536,647

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 3

; LENGTH: 277

; TYPE: PRT

; ORGANISM: Human

US-09-536-647-3

Alignment Scores:

Pred. No.: 4,37e-109

Score: 1325.00

Percent Similarity: 95.67%

Best Local Similarity: 91.70%

Query Match: 44.66%

DB: 4

Length: 277

Matches: 254

Conservative: 11

Mismatches: 12

Indels: 0

Gaps: 0

US-09-896-522-1 (1-1624) X US-09-536-647-3 (1-277)

```

QY      94 ATGCTTCCGCGGAGGCGAAGACTCCGAGACCCCGCGGAGGCGGACCGCTCCGAC 153
Db      1 MetAlaSerAlaGlyGlyGlySerGlnSerAlaAlaProGlnLysAspArgProGln 20
QY      154 CAGCGGCGCTCTGTTAAGGGGTGAGCGGCGGACCTCCAGCGGGAAGTCCGACCGTGTGT 213
Db      21 ProArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
QY      214 GAGAGATCATGAGTTGCTGGGACAGAACGAGGTGGAACGCGGACCGGAAAGTGTGTC 273
Db      41 GlnuyllelmetGlnleuLeuGlnlnAsnGlnValAspArgGlnArgLysLeuVal 60
QY      274 ATCTGAGCCAGGACGAGTTCTACAGAGTCTTGACCGGACAGGAAGGACGAGCTTG 333
Db      61 lLeuSerGlnAspArgPheTyrLysValAlaLeuThrAlaGlnGlnLysAlaLysAlaLeu 80
QY      334 AAAGACAGTACATATTTGACCATCCAGATGCTTTGATATATGATTTGATGACAGGACT 393
Db      81 LysGlnGlnTyrAsnPheAspPheProAspAlaPheAspAsnAspLeuMetHisArgThr 100
QY      394 CTGAAGACATCTGTGGAGGGGAAAACGGTGGAGGTGCCGACCTTATGATTTTGTGACAC 453
Db      101 LeuLysAsnIleValGlnGlnLysThrValGlnValProThrTyrAspPheValThrHis 120
QY      454 TCAAGTTTACAGAGACACGAGTGTGTCACCTGCGGACGAGTGTCTGTTTGAAGGACATC 513
Db      121 SerArgLeuProGlnThrThrValValTyrProAlaAspValAlaLeuPheGlnGlyIle 140
QY      514 TTGGTGTTCACAGACGAGAGATCCGGACATGTTCCACCTGCGCTCTTGTGACACCC 573
Db      141 LeuValPheTyrThrGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
QY      574 GACTCCGAGCTGAGGCTGTCTCGAAGAGTTCTCCGGACGTCGCCCGAGGAGGACCTG 633
Db      161 AspSerAspValArgLeuSerArgArgValLeuArgAspValGlnArgGlyArgAspLeu 180
QY      634 GAGCAGATTCTGACGACGATACACACCTTCGTGAAGCGGCGCTTCAGAGAGTTGCTGCTG 693
Db      181 GlnGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGlnGlnPheCysLeu 200
QY      694 CCGACAAAGAAGTATGCCGATGTGATCATCCACGAGAGTGGACAAATATGTTGCCATC 753
Db      201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
QY      754 AACCTGATCTGCGACGACATCCAGACATTTGAAATGTGACATCTGCAATATGGACCGA 813
Db      221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlnLysAspIleCysLysThrPheArg 240
QY      814 GAGAGGTCCAAATGGGCGGAGCTTCAAGCGGACCTTTCTGAGCGAGGAGCAACCTGGG 873
Db      241 GlnGlnSerAsnGlnLysSerTyrLysArgThrPheSerGlnProGlnLysAspHisProGly 260
QY      874 ATCTGACCTCTGGAACGAGTTCACATTTGAGTTCGAGACGAGACCGGAC 924
Db      261 ValLeuAlaThrGlyLysArgSerHisLeuGlnLysSerSerArgProHis 277

```

RESULT 3
US-09-134-001C-3618

; Sequence 3618, Application US/09134001C
; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

```

: SEQ ID NO 3618
: LENGTH: 216
: TYPE: PRT
: ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3618

Alignment Scores:
Pred. No.: 9.69e-32 Length: 216
Score: 451.00 Matches: 91
Percent Similarity: 63.68% Conservative: 44
Best Local Similarity: 42.92% Mismatches: 61
Query Match: 15.20% Indels: 16
Gaps: 4

US-09-896-522-1 (1-1624) x US-09-134-001C-3618 (1-216)
QY 166 CTGATAGGGGTGAGCGCGCGGCACTGCCAGCGGAAAGTGCAGCCGTGTGAGAAATCATG 225
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
D 15 lTelleGylIleAlaGlyGlySerGlySerIlyThrIlyThrIlyAlaAlaIleMet 34
QY 226 GAGTTGCTGGGACAGAAAGAGGTGGAAAGCGGCGAGCGGAAGGTGCATCTTGAGCAG 285
   ::|||
D 35 LysAsnLeu-----GluGlyHisSerValAlaIleuValGln 47
QY 286 GACAGGTTCTCAAG-----GTCTGACCGGACAGACAGAAAGGCCAGGCTTGAA 336
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
D 48 AspIlyTyTyTyLysAspGlnSerHisIleuThrPheGlnIlyuArgLysGluThr----- 65
QY 337 GGACAGATCAATTTTGACATCCAGATGCTTTGTATATGATATGATGATGACAGAGACTGTG 396
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
D 66 -----AsnIlyAspHisPProPheAlaPheAspAsnAspLeuIleHisAsnLeu 82
QY 397 AAGAACATCGTGAGAGGGCAAAACGGTGGAGGTGCGGACTTATGATTTTGTGACACACTCA 456
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
D 83 LysAspLeuArgAsnGlyLysProValGluValProThrTyArgTySerGlnHisThr 102
QY 457 AGGTTACCAAGAGACCGGTGTCTACCTCGGACCGAGCTGGTCTTGTAGCGGCATCTTG 516
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
D 103 ArgSerLysGluThrIleAlaPheAspProLysAspValIleIleValGluGlyIlePhe 122
QY 517 GTGTTTACAGCCAGAGATCGGGAGACATGTTTCCACCTCGGCGCTTCGTGGACACCGAC 576
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
D 123 AlaLeuGluAsnAsnThrIleuArgAspMetCysAspValIlyTyValAspThrAsp 142
QY 577 TCCGACGTCAAGGCTGTCTCGAAGGTTCTCCGGACGTGCGC---CGAGGAGGAGACTG 633
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
D 143 AlaAspLeuArgIleIleuArgArgLeuThrArgAspThrLysGluArgGlyArgThrMet 162
QY 634 GAGCAGATTTGACGCGACGATACCACTTCGTGAAGCCGGGCTTCGAGAGATTGCGCTG 693
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
D 163 GluSerValIleAsnGlnTyLeuAsnValAlaArgProMetHisGluGlnPheIleGlu 182
QY 694 CCGACAAAGAAAGTATGCCGATGATATATCCACAGAGAGTGCACATATAGGTGGCATC 753
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
D 183 ProThrLysLysHisAlaAspIleIleIleProGluGlyLysSerAsnLysValAlaIle 202
QY 754 AACCTGATCGGACGACATCCAGAGACTTTGAT 789
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
D 203 AspIleMetThrThrLysIleGlnSerLeuValSer 214

RESULT 4
US-09-107-532A-5024
: Sequence 5024, Application US/09107532A
: Patent No. 6583275
: GENERAL INFORMATION:
: APPLICANT: Lymn A Doucette-Stamm and David Bush
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
: NUMBER OF SEQUENCES: 7310
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: GENOME THERAPEUTICS CORPORATION
: STREET: 100 Beaver Street
: CITY: Waltham

```

```

1 STATE: Massachusetts
2 COUNTRY: USA
3 ZIP: 02354
4
5 COMPUTER READABLE FORM:
6 MEDIUM TYPE: CD/ROM ISO9660
7 COMPUTER: PC
8 OPERATING SYSTEM: <Unknown>
9 SOFTWARE: ASCII
10
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/09/107,532A
13 FILING DATE: 30-Jun-1998
14
15 PRIOR APPLICATION DATA:
16 APPLICATION NUMBER: 60/085,598
17 FILING DATE: 14 May 1998
18 APPLICATION NUMBER: 60/051571
19 FILING DATE: July 2, 1997
20
21 ATTORNEY/AGENT INFORMATION:
22 NAME: Ariniello, Pamela Deneke
23 REGISTRATION NUMBER: 40,489
24 REFERENCE/DOCKET NUMBER: GTC-012
25
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: (781)893-5007
28 TELEFAX: (781)893-8277
29
30 INFORMATION FOR SEQ ID NO: 5024:
31
32 SEQUENCE CHARACTERISTICS:
33 LENGTH: 210 amino acids
34 TYPE: amino acid
35 TOPOLOGY: linear
36 MOLECULE TYPE: protein
37 HYPOTHEetical: YES
38 ORIGINAL SOURCE:
39 ORGANISM: Enterococcus faecium
40
41 FEATURE:
42 NAME/KEY: misc feature
43 LOCATION: (B) LOCATION 1..210
44
45 SEQUENCE DESCRIPTION: SEQ ID NO: 5024:
46
47 US-09-107-532A-5024
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
9
```

```
Db 114 GlyIleuLeuLeuGluAspGluArgLeuArgSerLeuMetAspIleValTyrVal 133
|||
QY 568 GACACCGCATCCGACGTGAGCTGTCTGAAAGATTCTCGGAGCGTG---CGCCGAGG 624
|||
Db 134 AspThrAspAspAspIleArgIleIleArgArgIleValArgSerMetGluValArgGly 153
|||
QY 625 AGGAGCTGGAGCAGATTCTGACCGACGATCCACCTTCGGAAGCCGGCTTCGAGAG 684
|||
Db 154 ArgThrLeuAspSerValIleGluGlnPyrLeuThrValValValValValValValVal 173
|||
QY 685 TTTCGCTGCGCAGCAAGAGATAGCCGATGTGATCATCCACAGAGAGTGACATATG 744
|||
Db 174 PheIleGluProThrValArgTyrLeuAspIleIleValProGluGluGluValAsnHis 193
|||
QY 745 GTTGCCATCAACGTGATGTGACACATCCAGACATTTGAT 789
|||
Db 194 ValAlaIleAspLeuIleThrThrValAlaValAspPheLeuAsn 208
|||

RESULT 5
US-09-198-452A-783
; Sequence 783, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: and thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198-452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 783
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-783

Alignment Scores:
Pred. No.: 2,79e-28 Length: 222
Score: 412.00 Matches: 84
Percent Similarity: 63.01% Conservative: 54
Best Local Similarity: 38.36% Mismatches: 59
Query Match: 13.89% Indels: 22
DB: Gaps: 6

US-09-896-522-1 (1-1624) x US-09-198-452A-783 (1-222)
QY 166 CTGATAGGGGTGAGCGGCGACATGCGAGCGGAGAGTGCACCGTGTGAGGAAGATCATG 225
|||
Db 8 IleIleGlyIleThrGlySerGlyAlaGlyIleThrThrLeuGlnAsnIleVal 27
|||
QY 226 GAGTGTGCGGAGCAAGAGAGTGAACAAGCGGAGCGGAGAGTGTGATCATCTGAGGCA 285
|||
Db 28 GluIlePheGlyGluAsp-----ValSerValIleCysGln 39
|||
QY 286 GACAGGTTCTCAAG-----GTCTGACGGCAGAGCAAGAGCGCAAGGCTTGAAGA 336
|||
Db 40 AspAsnIleThrValAspArgSerHisIleTyrThrProGluValArgAlaAsnLeuIle--- 58
|||
QY 337 GGAAGATCAATTTGACATCCAGATGCTTGTATATGATTGTGATGACAGAGACTCTG 396
|||
Db 59 -----TTPAspHisProAspAlaPheAspAsnAspLeuIleSerAspIle 74
|||
QY 397 AAGAATCTGTGAGGGCAAAACGTTGAGGTCGCGCATATGATTTTGTGACACTCA 456
|||
Db 75 LysArgLeuLysAsnAsnGluIleValGlnAlaProValPheAspPheValLeuGlyAsn 94
|||
QY 457 AGG---TTACAGAGACACAGCGTGTCTACACCTGCGAGCGTGTCTGTTGAGGAGCATC 513
|||
Db 95 ArgSerLysThrGluIleGluThrIleTyrProSerLysValIleLeuValGluGlyIle 114
|||
QY 514 TTGGTGTCTTACAGCCAGAGATCCGGGACATGTTTCCACCTGCGCTTCTGTGAGACCC 573
|||
```

```
Db 115 LeuValPheGluAsnGlnGlnIleuValArgAspLeuMetAspIleArgIlePheValAspThr 134
|||
QY 574 GACTCCGACGTCAAGCTGTCTCGAAGAGATTCTCCGGACGTGCGC---CGAGGAGGAGC 630
|||
Db 135 AspAlaAspGluArgIleLeuValArgSerMetValArgAspValGlnGlnIleGlnIleAspSer 154
|||
QY 631 CTGGACAGATTCTGACCGAGATACACACCTTCGTGAACCGGCTTCGAGAGCTTCG 690
|||
Db 155 ValAspCysIleMetSerArgTyrLeuSerMetValLysProMetHisGluLysPheIle 174
|||
QY 691 CTCGCCACAAGAGATATCCGATGTGATCATCCACGAGAGTGGACATATGTTGCC 750
|||
Db 175 GluProThrArgLysThrAlaAspIleIleValHisGlyAsnTyrArgIleAsnValVal 194
|||
QY 751 ATGACCTGATCGTCGAG-----CACATCCAGACATTTGATGATGATC 795
|||
Db 195 ThrAsnIleuSerGlnLysIleLysAsnHisIleGluAsnAlaLeuGluSerAsp 213
|||

RESULT 6
US-09-252-991A-30843
; Sequence 30843, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196-136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30843
; LENGTH: 663
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30843

Alignment Scores:
Pred. No.: 1,07e-09 Length: 663
Score: 204.50 Matches: 147
Percent Similarity: 29.20% Conservative: 46
Best Local Similarity: 22.24% Mismatches: 183
Query Match: 6.94% Indels: 285
DB: Gaps: 37

US-09-896-522-1 (1-1624) x US-09-252-991A-30843 (1-663)
QY 1619 CCGCGCAAGCCTCG-----GTGTGATGCTGGGTGCTTCCAGGGCAT 1578
|||
Db 132 ProAlaSerAlaIlePyrIlePheValProAlaArgCysCysIlePyr-----ProAlaPro 149
|||
QY 1577 CGGCTCGCTTCCTGCGCTTTAAGCACAAGAGGGGAAAC-----1533
|||
Db 150 ProAlaProAlaSerAlaGlyArgAlaCysCysAlaSerProAsnArgArgArgGlu 169
|||
QY 1532 -----ATCCCT 1527
|||
Db 170 ProTyrProProSerProTyrAlaSerArgAlaGlyProAlaSerCysGlyArgProPro 189
|||
QY 1526 CAGTGGCTCCCATATCCGTGAGGACAGCTGCAATCGAACCCTGGGTCAATTAAT 1467
|||
Db 190 AlaCysSerProValAlaThrAlaProThrAlaIleCysSerProProSerAlaArgSer 209
|||
QY 1466 AGTGTGTAGGCACAACAGCTGTGTATGAAACAGAAAACATGAGG---1410
|||
Db 210 AlaTyr-----LysProCysAlaCysAlaIleValAlaGlySerAlaArgSer 225
|||
QY 1409 -----ACATGTAACATGTAAAAAGG 1389
|||
```

D	b	226	ProAlaAenAlaTrpProAlaAlaSerAlaTrpCySProThCysAenSerAlaArgPro	245
Q	y	1388	AAAGACTGTTCACAAAATTGTAAGTAAAGTGTGAAGAACGAGATCAGA	1322
D	b	246	SerAlaIleAlaSer-----	250
Q	y	1328	CTGAAAAAACTCTCTCCCACTGTGGGTCACTGTCAACAA-----ACATCAGGC	1277
D	b	251	-----ThAlaArgArgSerProSerThrsAngly	260
Q	y	1277	CAGCCAGTGTCTAGGCTGTCTCCAAATTCCCAATAATGCGCTCAATCTCCACAG	1211
D	b	261	ArgPro-----AlaAlaCysProThrThrx	269
Q	y	1217	AAGCTCCCAAGCTCTTCGACATTTCTGTGGCATTTCTCAGTACCTAAGAGGATCTTAA	1158
D	b	270	ProProAlaAenAlaProAlaSerAlaAlaAla-----	280
Q	y	1157	AACCGCAACGAGCTTAAGTGGCTGAAAACCTGA---GGAACCCGTGTCAGTGCCAGA	1107
D	b	281	-----AenTrpProTrpTrpLysAlaProSerProAlaMetProProSerThrsAla	298
Q	y	1100	AGTTGAGTGTAGTACATGATCGAGTTTC-----CACTCGTAGTGAAGAAAGGCTTG	104
D	b	289	-----ThrsArgLeuAlaAlaCysGlyHisAsp--GlyAspAlaGlyArg	313
Q	y	1046	CTGCTAACACTCCCTCG-----GGGTGCGCCGAG---AGGA	101
D	b	313	LasErGlThnTrProTrpArArgAspThrsProCysAlaProAnCySValaAntTrpArg	333
Q	y	1013	AGCACTGGGTGTGGTGGCGCTCCCAAGCTCAGTCCCTGAACAACATCGCGGCGGGA	954
D	b	333	rgArgIleProTrpSerAlaValAlaAlaLysThrsProTrpProArgTrpPro---GlyM	352
Q	y	953	GACCTGCCTAGAGGCTGGCAGCCCTCAAGGCGTGTCTGTCTGTGACATCCCAATGTGAC	894
D	b	352	etProAla--GlyProArgProSerCysThrsAlaValAlaAlaAlaTrpAlaPro-----	368
Q	y	893	CGTTGCCAGAGGTACAGATCCCAAGGTGTGCTCCGTGCTCGAAGAAAGGTCCGCTTGA	834
D	b	369	-----AlaArgGlyClyLysTrpAlaAlaAlaProArgAnSerCysSe	382
Q	y	833	CTCGGCCATTGGACCTTCTCTCGGTGCATTTTGACAGATGTACCATTCAGAAATGTCTGG	774
D	b	382	rAlaGlyArg-----LeuArgArgArgValArgArg-----	392
Q	y	773	ATGTCTGCACGATCAGGTGTAATGGCAACCATATTGTCCACTCTCTGTGGATGATCACA	714
D	b	393	-----SerAlaArgArgAla--TrpArgProTrp-----	401
Q	y	713	TCGGCATCTTCTTGTGGCAGAGAGA---ACTCCT-----CGAAGGCCGCG	669
D	b	402	-----ProAlaGlyArgAlaThrsProAlaSerAlaAlaArgAspSe	416
Q	y	668	TTCACGAAGG-----TGTGTACTGTCCGT	645
D	b	416	rAlaSerArgGlyArgProAlaThrsAlaAlaAspHisProAlaAlaTrpValSerAlaAl	436
Q	y	644	AGAA-----TGTGTCTCAAGTCCCTCCCTC-----	620
D	b	436	argArgThrsSerSerAlaProIleAlaGlySerAlaProGlySerGlyThrsAlaProAr	456
Q	y	619	GGCGACAGTCCCGAGAACTTTCGACAGACGCTGAC	582
D	b	456	gCyHisArgValaArgLysAspGlyAlaGlyProAlaSerThrsGlyArgSerArgArgAr	476
Q	y	581	TCGAGTCCGTGTCCA-----CGAAGAGCGCAGGTGAAC	546
D	b	476	gTrpAlaAlaAlaProAlaArgArgAlaArgThrsGlyIleArgArgThrsAlaAlaAlaAl	496
Q	y	545	ATGT-----CCCGATCTCTCTGGCTGTGAACACCAAGATGCCCTCAACAGAACACG	492
D	b	496	aCysTrpProProAlaProGlyArgArgThrsProAlaAlaArgArg-----	512

```

QY 491 TCCGAGGGGTAGACCAACCGGTGCTCTGTGTAACTTGAAGTGTGTACAAATAATCATAGGTC 432
Db -----
QY 513 -----ArgSer1aIargTh 517
431 GGACACCTCCACCGTTTGGCCCTCCACAGATGTTCTTCAGAGTCCTGGCATCAAAATCATTA 372
Db :||||| ||| ||||| ||| |||
517 TAlaProAlaProAlaIleProAlaIleArgProSerAlaSerAlaThnGly----- 534
QY 371 TCAAAAGCATGTGATGGTCAAAATTTGACTGTCTCTTCAAGAGCCTTGGCTTGTCTGTCT 312
Db -----TriPro----- 536
QY 311 GCCGTGAGACCTTGTAGAACCTGTCT--GGCTCAGAGTAGACACCTTCCGCTGCCG 255
Db :||||| ||||||| :||| ||||||| |||
537 -----GlyProAlaLeuAlaCysProAlaIleAlaIleGlyArgArgProSerProAlaPr 554
QY 254 TGTTCGA-----CCTCGTTCTGTGCCACAGAACTCATATCTTCTCACAACAGGTC 204
Db :||| ||| ||| :||| |||
554 oAlaProAlaProAlaProAlaIleAlaIleArgProAlaIlePro----- 567
QY 203 GACTTCCCGCTGGCAGTCCGCGCGCTCACCCCTATCAGGAAGGCGCGTGTGGGACGG 144
Db -----ArgArgAlaPro-----G1 572
568 -----
QY 143 TCGGCTCCGCGCGCGGGCTCTCCGAGATCTTTCGCTCCCG-----CGAGACCATCTCG 90
Db :||| ||||||| :||| ||| |||:|
572 YArgSerProAlaThnAlaProAlaIleAlaIleAlaIleAlaIleGlyProThnProAlaIaR 592
QY 89 GCTTCGCTCCGCGCGCATCGAGTCCCG--CGCCGCGCCCTTCCCGGGGCGCGCGCGC 33
Db :||||| ||| ||| ||||||| ||| |||
592 gProProAlaIleAlaIleThnGlyArgProSerAlaIleProAlaIleGlyProAlaIleArgArgGlyA 612
QY 32 C-----CGCCACGAGCGGAGGTGCG 14
Db -----
612 lArgTrpProTrpArgSerAlaIleArgCysSerProAlaIleArgTrpAlaIleArgSer 630

```

```

RESULT 7
US-09-252-991A-25394
; Sequence 25394, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ. ID NOS: 33142
; SEQ. ID NO 25394
;
; LENGTH: 369
;
; TYPE: PRT
;
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25394

```

Alignment Scores:	
Pred. No.:	1.12e-09
Score:	203.00
Percent Similarity:	35.76%
Best Local Similarity:	27.64%
Query Match:	6.89%
DB:	4
	7
	Gaps:
	14
	27
	36
	Length:
	13
	Matches:
	Conservative:
	37
	Mismatches:
	15
	Indels:
	17
	Gaps:
	14

```

US-09-896-522-1 (1-1624) x US-09-252-991A-25394 (1-369)

OY      1284  ATCAGCGAGCAGTATC-----TAGCGTCTCTCATTTCCCAATATG 1237
      :::|||||:::      |||  |||||:::|||||:::

Db      4  LeuArgProAlaAlaPhePheProLeuProG|ValAlaProSerIleAlaProLeuSer 23

```

```

QY 1236 TGCCCTCATCTCCACAGAAAGCCCTCCAGGCTTCCTCCACATTCCTGTGGCATTTCTCAG 1177
Db      :::::
QY 24 GlyAlaArgThrMetProCysGlySerArgAlaGlyArgGlyLysArgAlaArgHis 43
Db      :::::
QY 1176 TGACCTAAGAGGATTTAAACCGCAGACGACCTTAAGTGGCTGAAAACCTCAGAGACGCC 1117
Db      :::::
QY 44 CysProAlaArgProAlaArgProSerArgSerAlaValAlaGlySer--ArgAlaAla 62
Db      :::::
QY 1116 TGTCAGTGTCCCGCAGAGTTGAGTGTGAGTGACACATCTGAGTTTCCACTCCTGAGTGAG 1057
Db      :::::
QY 63 SerArgCysProSer----- 67
Db      :::::
QY 1056 GAAGGCTCGCTGCTAACACTCCCTGGAGTCC-----GCCAGAGAGAACAGCTGG 1006
Db      :::::
QY 68 -----AlaAlaThrPro-----CysHisProThrLeuArgArgArgArgTyr 81
Db      :::::
QY 1005 GTGTGGTGGGGGCGCCAGGCTCAGTCCCTGAGACACACATGCGGGGGGAGACCTGCC 946
Db      :::::
QY 82 --ProGlySerArgProProArgSerSer-----ThrGlyProGlyArgArg-----P 97
Db      :::::
QY 945 CTGAGGCTCGGACGCGCCCTCAGTGGGGGTCTGCTGCTGAGCTCCAAATGTGACCGTTGGC 886
Db      :::::
QY 97 roProArgArgArgProProArgArgSerAlaAlaAlaSerProArg-----H 113
Db      :::::
QY 885 AGAGGTGACATCCCGAGGTGGTCCCTGCTCAGAAAAGGTCCGCTTGAGTCCGCC 826
Db      :::::
QY 113 iSAgtThrSerAlaProArgGlyPro-----GlyArgArgThrProAlaSerSer----- 129
Db      :::::
QY 825 ATTGAGCCCTCTCGGTGCCATTGTCAGATGTACACATTCAGATGTCCTGAGTGTG 766
Db      :::::
QY 130 -----SerGluArgProGlyThrAla- 136
Db      :::::
QY 765 CACGATCAGTTGATGCAACCATTTGTCCACTCCTGCGAGATGATCATCGGCATA 706
Db      :::::
QY 137 -----ProAlaProPro-----ArgArgT 143
Db      :::::
QY 705 CT-----TCTTGTGGCGAGGAGAACTCTCGAAGCGCG 670
Db      :::::
QY 143 hrAlaProGlyTyrProGlyProSerAlaAlaGlyAlaGlyArgProAlaArgArgProAla 163
Db      :::::
QY 669 CT-----TCAGAGGTGGTGTACTGGCTGAGATTCGCTCCAGGTCCCTCCCTCG 619
Db      :::::
QY 163 laArgPheProProProArgArgCysArgIle--ProAlaGlyProGlyProPro--- 180
Db      :::::
QY 618 GCGCAGCTCCCGGAGAACTTTCGAG--ACAGCTGACGTGCGAGTCCGAGT----- 569
Db      :::::
QY 181 --ArgArgProGlyArgProTyrArgGlyThrAla-----ArgTyrArgGlyArgGlyC 198
Db      :::::
QY 568 -----CCAGAAAGGCGCAGGTGGAACATGTCGCGATCTTCGCTGTA 523
Db      :::::
QY 198 ysCysAlaProAlaProArgArgArgSerProAlaIleProAspAlaArgProAlaCysA 218
Db      :::::
QY 522 GAACACCAAGATGCCCTCAACAGAACAGACGTCGCGAGGTAGA----- 479
Db      :::::
QY 218 laAlaProArgAlaGlyAlaGlyAlaAlaArgProAlaGlyGlnGlySerProAlaArgC 238
Db      :::::
QY 478 -----CCACCGTGTCTCTGTAACCTTGAGTGTGACAAATCATATGAG--TCGG 430
Db      :::::
QY 238 ysAlaAlaProSerArgProAlaAlaAlaArgSerAlaThrThrGlyArgArgProSerA 258
Db      :::::
QY 429 CACCTCCACCGTTTGGCTCCACGAGTGTCTTCAAGTCTGTGATCCTGATCAATATTATC 370
Db      :::::
QY 258 laSerArgAlaGlyArgProAlaArgCysProAlaGluArgArgAlaThr----- 274
Db      :::::
QY 369 AAAAGCATCTGAGTGTCAAAATTGTACTGCTTTCAAGGCTTGGCTTCTGCTGC 310
Db      :::::
QY 275 -----AlaProAlaAlaHisArgAlaTyrProProProAlaP 287
Db      :::::
QY 309 CGTCAGACCTTTGAGAACTGTCTGCTGCTGAGATGACACCTCCGCTCCGCTGTC 250
Db      :::::
QY 287 roAla-----ProAlaAlaProAlaAlaP 295
Db      :::::
QY 249 CA-----CTCTGTTCTGTCCAGCACTCATATCTTTCACACAGCGTGACTTCC 196

```

```

Db      :::::
QY 295 roGlyAlaAlaArgSerArgProAlaGlyPro-----ArgArgGlyArgArgA 311
Db      :::::
QY 195 GCTGGCAGTGTCCCGCGCTGACCCCTATCAGAAAGGCGCGCTGGTGGCAGACGTCGCTC 136
Db      :::::
QY 311 rgaArgArgCysArgSerAlaPro-----AlaGlySerGlyGlyAsnPro- 325
Db      :::::
QY 135 CGGCGCGGGGCTCTCGCAGTCTTCGCTCCCGCGCAAGCCATCTCGAGCTCCGCTCCGC 76
Db      :::::
QY 326 ArgLysArgArgLysGlyCysCysArgArgArgArgSer-----AlaAlaPro 341
Db      :::::
QY 75 GCATCGGGTCCCGCGCGCGCTTCCTCCCGGCGCGCGCGCGCGCGCGCGAGT 16
Db      :::::
QY 342 AlAlaAlaProGlyAla-AlaAlaSerProGlySerGlySerProProAlaAlaArgPr 361
Db      :::::
QY 15 CGGAGCGGACCCC 3
Db      :::::
QY 361 oValAlaAsnPro 365

```

RESULT 8

```

US-09-252-991A-23264
: Sequence 23264, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ. ID NOS: 33142
: SEQ ID NO 23264
: LENGTH: 568
: TYPE: PRT
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23264

```

```

Alignment Scores:
Pred. No.: 5,67e-09 Length: 568
Score: 196.00 Matches: 119
Percent Similarity: 34.73% Conservative: 30
Best Local Similarity: 27.74% Mismatches: 150
Query Match: 6.65% Indels: 130
DB: 4 Gaps: 23

```

```

US-09-896-522-1 (1-1624) x US-09-252-991A-23264 (1-568)

```

```

QY 1127 TCAGGAACGCTGTGTCAGTGTCCAGCAAGTGTGAGTGCACATCTGATTCAC 1068
Db      :::::
QY 9 SerGlySerProAlaThrCysArgSerThr----- 18
Db      :::::
QY 1067 TCCTGATGAGAGAGGCTGCTGTAACACTCCCTGGGGGTGCGCGAGAGAGACAGT 1008
Db      :::::
QY 19 -----ThrArgSerSerThrMetProCysAlaTyrProAlaPro---ProAlaThr 34
Db      :::::
QY 1007 GGGTGGGGGTGCGCTCCAGGCTCAGTCCCTGAACACACATGCGGGGCGGAGACTTG 948
Db      :::::
QY 35 GlyArgAlaTyrPro-----SerAlaProAspLeuAlaArgArgSer----- 48
Db      :::::
QY 947 CCCTGAGGCTCGGCGCCCTCAGTGGGGGTGCTGC-----TGG----- 908
Db      :::::
QY 49 ProTyrSerValAlaAlaGlyArg-GlyArgCysCysAlaAlaAlaSerSerTyrCysAr 68
Db      :::::
QY 907 -ACTCCAAATGTGACCGTTTGCAGAGTCCAGATCCAGGAGTGTCCCTGCTCAGAA 849
Db      :::::
QY 68 gthrThrSerArgAlaValProTyrArgSerCysAlaIleGly----- 82
Db      :::::
QY 848 AAGTCGCTTGTGAGTCCGCGCATTTGAGACCTCTCGGTGCGCATTTGCAGATGTACCA 789
Db      :::::

```

```

Db      83  ----SerGlyCysLeuArgSer---TrpThrSerArgGly-----CysArgTrpThrAr 98
Qy      788  TTGAGATGTCCTGGAGTGTCTGCAGATCAGTGTGATGGCAACCAATATGTCCACCT 729
Db      98  gser---CysAsnSerCysSerThrArg-----TyrArgProPro-- 110
Qy      728  CGTGGATGATGATCGGATACCTTTGTGCGGAGAGAACTCCTGAGAGCCGCG 669
Db      111  -----AlaTyrGlnAlaIleAlaArgProAlaGlyAlaValaArgArgAlaProAl 127
Qy      668  TTCACGAGAGGTGTGTACTGCTGCAGAACTGTCTCCAGGTCCTCCCTCGCGCAGCTCC 609
Db      127  aalaglyHis-----ArgProGlyGlnProValaGlyAlaArgPr 140
Qy      608  CGGAGAACTTTGTGAGACACCTGACCTCGGAGTCCGAGTCCACGAGAGCCGAGGTGG 549
Db      140  oalavalAlaCysAlaThrAlaTyrIleLeuValaGlyAlaArgArgAlaAlaProAl 160
Qy      548  AACATGTCCTGGATCTCTGCTGTGTAGAACACCAAGATGCTTCAACAGAACACGATCC 489
Db      160  aargProAlaGlyArgProValaGlyAlaProProGlyGlyAlaGlnAlaGlnProProAl 180
Qy      488  GCAAGGTAGACCACTGCTGTCTGTGTA-----ACCTTGAAGTGTGTACAAATCATAG 435
Db      180  aglnArgGlnProAlaAlaGlyAlaLeuGlnArgSerAlaSerAlaSerAlaAlaHisIle 200
Qy      434  GTGGGCA-----CCTCACCGTTTGGCCTCCACGATGTTCTTCAAGTCTGTGC 384
Db      200  uSerAlaSerGlySerProAspProValaArgSerPro-----Al 213
Qy      383  ATAAATCATATTATCAAAAGCATGTCAGTGTGTCAAAATTGACTGTCTTCAAGGCTTG 324
Db      213  aglyArgHisGlyPheArg----- 219
Qy      323  GCCTTGTCTGTGCTGCTGCTGACGACCTGTGACGCTGTCTGCTGCTGCTGCTGCTGCTG 273
Db      220  -----ValProAlaAlaGlyGlyArgThrAlaProGlyArgGlyHisArgGlyArg 236
Qy      272  ACCAGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 239
Db      236  glnProGlyAlaAlaAla---GlnProArgSerLeuArgValaSerProLeuArgAlaGly 255
Qy      238  -----GTCCAGCAACTCCATGATCTTCTTCACACAGCGTGCAC 201
Db      255  yThrAlaGlnProAlaAlaProValaAlaAlaAlaProLeuProGlySerThrArgArg 275
Qy      200  TTCCTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 141
Db      275  pProArgLeuProGlyLeuArgProProLeuArgArgArgThrAspGly-----GlyArg 293
Qy      140  GCCTCGCGGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 90
Db      293  gProLeuAlaGlnProAspArgArgArgThrArgArgProAlaProArgProArgProGly 313
Qy      89  GCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 47
Db      313  lProProValaAlaArgIleProArgArgArgHisProAlaProAspArgLeuGlnGlyHis 333
Qy      46  -----CGGAGCCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 27
Db      333  sArgSerGlnThrHisAlaAspArgProArgIleProGlyArgThrArgProAlaAspPr 353
Qy      26  AGGCGCGAGGTCCGAGGCGGAGCC 4
Db      353  oleuPro-AlaArgLeuArgPro 360

```

```

RESULT 9
US-09-252-991A-31116
; Sequence 31116 Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: MARC J. RUBENFIELD et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

```

```

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31116
; LENGTH: 664
; TYPE: PR
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31116

Alignment Scores:
Pred. No.: 9.13e-09 Length: 664
Score: 194.00 Matches: 124
Percent Similarity: 29.39% Conservative: 25
Best Local Similarity: 24.46% Mismatches: 162
Query Match: 6.54% Indels: 196
DB: Gaps: 23

US-09-896-522-1 (1-1624) x US-09-252-991A-31116 (1-664)
Qy      20  CGGCGCTGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 70
Db      240  ArgArgAlaValaAlaAspProAlaGlyArgProGlyGlnGlyHisGlyArgArgProArg 259
Qy      71  GATCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 130
Db      260  AspHisArgPheGlnArg-----ArgArgArgAlaAspProArgProArgArgArg 275
Qy      131  CGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 190
Db      276  ArgArg-GlyArgProGlyGlnGlyAlaAla-----HisArgArgGly 289
Qy      191  CCAAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 230
Db      289  yHisArgGlnArgProArgSerGlyAlaArgHisGlyValaHisValaAspArgArgGlnGly 309
Qy      231  -GCTGGAGCAGAACAGAGTGTGAACAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 289
Db      309  nProGlySerGlyProGlyGlyArgProAlaProGlyGlyArgGlnProAspProAlaGly 329
Qy      290  GGTTCACAAAGT-----CCTGACGCGAGCAGAAAGCGCA 325
Db      329  nGlyGlnGlnGlyAspHisArgLeuArgProHisArgProGlyGlnGlyArgGlyHis 349
Qy      326  AGGCTTGAAGAAGCAGTCAATTTTGACCAATCCAGATGCTTGATGATGATTTGATGC 385
Db      349  sGly----- 350
Qy      386  ACAGGACTGTGAAGCATGTGAAGGCAACAGGTGAGGTGCGGACCTGATGATTTG 445
Db      351  -----GlnGlnGlnProGlyGlyArgArgGlyAlaGly----- 361
Qy      446  TGAACACTCAAGAGTTTACAGAGACCGGTGATCTTACCTGCGGAGCGTGTCTGTTTG 505
Db      362  -----AspArgGlyAlaLeuProValaProArgGlnHisProArg 374
Qy      506  AGG-----CATCTT----- 515
Db      374  gGlyValaAspHisArgHisHisArgProAlaValaAspAlaValaHisLeuHisArgArg 394
Qy      516  -----GGTGTTCACAGCAGAGATCCGAGACATGTTTCCACCTGCGGCTTCTGCG 565
Db      394  yArgGlnGlnGlnGlnArgGlnProAspGlnProArgArg-----AlaGlyLeuArg 411
Qy      566  TGAACACGACTCCGAGCTGAGCTGTCTGCAAGAGT-----TCTCGGAGCGTGGCGGCG 622
Db      411  gHisHisArgArgArgArgArgGlyAspArgArgGlnArgHisArgProProProGlyProArg 431

```

```

Oy      623  GGAGGGA-----
Db      411  gProGlyProProAlaProProAlaAspProAlaGlyThrPheProAlaGlyLeuAlaGAr 451
Oy      630  -----CCTGGAGCAGATTCTGACCGCAGTACACACCTTCGTGAAGCCGG 673
Db      451  gValGInGlyAlaProGlyAlaGlyLeuAlaGProAspHisHis----- 466
Oy      674  CCTGGAGAGATTGCGCTCCGCGACAAAGAAAGTATGCCGATGTATCTATCCACAGAGAG 733
Db      467  ---HisGlyGlyValProAlaAspLeuArg-----ProHisArgGlyAlaGGL 481
Oy      734  TGGACAAATAT-----GGTTGCCATCAACCTGATGTCGACAGACATCCAGCATTC 784
Db      481  yGluAspValProProAspGlyValHisArgGlyHisArgHisAlaAlaArg-ArgAspAsp 501
Oy      785  TGAATGCGACATC-----TGCMAATGGCACCCGAGAGCGTCCATG 826
Db      501  rGleuGlyAspLeuArgProGlyGlyAspArgAlaValAlaHisArgGlnGlyGlnGly 521
Oy      827  GCGCGAGCTAC----- 837
Db      521  LyGlyGlyLeuAlaArgHisAlaProAlaGlyAlaSerGlyLeuAlaThrGlyAlaAlaLeuGly 541
Oy      838  -----AAGCGACCTTTCTGAGCCAGGAGCACACCTGT 871
Db      541  lArgAlaProAlaGlyAlaGlyAlaArgAlaArgProAlaArgAspProAlaHisArgAlaG 561
Oy      872  GGATGCTGACCTCGGCAACCGTCACTTTGGAGTCCAGACAGACAGCCACGTGAGGG 931
Db      561  LyGlyLeuAlaAspGlyGlnArgValHis-----SerGlnProGlnAlaArgGly 577
Oy      932  CTGCGGACCTCAGGCGAGGCTCTCCCGCCGCGACATGTGTTCAGGAGCTAGCGTGGGG 991
Db      577  rGlyLeuArgHisAlaGlyProAlaGlyAlaGlyHis-----GlnProAsp- 591
Oy      992  ACGCCCAACCCACACCCACTGCTCTCTCGCGCGCACCCAGGGAGTGTTCAGCCAGG 1051
Db      592  -----ProValGlyArgAspAlaAlaAspProGlyGlnGlyAla 604
Oy      1052  CCTTCCTCATCAGAGATGAGAACTCAGATGTCATCAGACACTCACTCTGTCGACAC 1111
Db      604  la-----AspGlyGlnValProGlyAspA 612
Oy      1112  TGACAGCGCTT-----CCTGAGGTTTTCAGCC 1138
Db      612  rGalArgLeuAlaArgProHisArgAspArgAlaArgAspArgLeuArgProAspAlaAlaGlu 632
Oy      1139  ACTTAGGCTCGTGGCGGTTTAAAGATCCCTCAGTGCTACCTGAGAAATGCCACAAATGTG 1198
Db      632  rGlyLeuArgGlnLeuArgAspAlaGlnAlaAlaGluProValAlaGlyProGlnGlyValAla 652
Oy      1199  CAGGAAGCCTGAGAGC 1215
Db      652  laGlyGlyAlaAlaGly 657

RESULT 10
US-09-252-991A-32424
Sequence 32424, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT FILING DATE: US/09/252.991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 32424

```

	:	LENGTH:	341
	:	TYPE: PRT	
	:	ORGANISM:	Pseudomonas aeruginosa
	:	US-09-252-991A-32424	
 Alignment Scores:			
Pred. No.:	2,09e-08	Length:	341
Score:	188.50	Matches:	103
Percent Similarity:	31.62%	Conservative:	26
Best Local Similarity:	25.25%	Mismatches:	139
Query Match:	6.40%	Indels:	140
DB:	4	Gaps:	25
 US-09-896-522-1 (1-1624) x US-09-252-991A-32424 (1-341)			
QY	1022	CCGAGAGA-----AGCAGTGGGTGTGGGTGGCGTCGCCAGGCTCAATGCCGAACAC	965
Dd	2	ProArgAsnAlaCystrHisSerIalAcysArgrITPProProProglYSerIalAla-----	19
QY	968	ACATGCCGGCGGAGAACATGCCCTCGAGGCTCGGCAGCCCCCTCATGAGGGGTCTGCTGTG	909
		::::	
Dd	20	-----MetProglYgLy-----AlaPro-----	25
QY	908	GACTCCAATGTGACCGTTTGCCAAAGTCAAGCATC--CCAGGGTGTCCCTTGCTCA	852
		:::	:::
Dd	26	--ThrIalLeSerYsAlaSerSerIalAlaIleSerProSerThnSerProglYAla	44
QY	851	GAAGAAGTCCGCTTAGTACGTCCGCCCATTTGACACCCTCGGGTCCATTGGCAATGTCA	792
Dd	45	ProLys-ProProThnSerThnSerCySphenIalArgThnSerIalAlaCys-----	61
QY	791	CCATTGCAATGTCTTGATGTGCTGCAAGATCAGATTGATG--CAACCATATTGTCC	735
Dd	62	-HisTrpThnCysTrpArgProAlaProAlaSerProAlarIpIleProProTYrCys--	80
QY	734	ACTCTCTGTGGAGTATCATCATTGCCGATACCTTTTGTGGCAGGACAGAACTCCTCGAAG	675
Dd	81	-----LeuProArgSerProAlaTy	87
QY	674	GCCGGCTTCACGAAGGTGGTGTACTGCGCAGAAATGTCTCCAGATCCCCTCGGCGC	615
Dd	87	sProArgGlYSerIalATrPCySArXArgSer-----ProProArgserIl	102
QY	614	ACGT-----CCCGAGAACTCTTGACAGACCTGACCTGACGTGGTGTCC	567
Dd	102	eArgProThnTrpTrpProglYSerCys-----AsnArgCys--	114
QY	566	ACGAAGAGGGCGAGGGAACATGTCCC-----GGATCTCCGTGGCTGGAACA--	518
Dd	115	---ThnglYSerIalThnValAlaProAlaglyThnSerSerProArgCysAlaIaln	133
QY	517	-----CCAAAGTGCCTCAACAGAACCA-----CGTCGCGAGG	483
Dd	133	rArgThnSerIalATrPThnArgCySPrometArgSerSerIalThnrgAlaProProse	153
QY	482	TAAACCAACCGTGG-----TCTTGGAACCTTGAAGTGTGTACA	444
Dd	153	rSerProglYTPrCYeThnThrCYeglyThnAlaSerProAlarghiIsSerITPserTh	173
QY	443	AAATCATAGGTGGCACCTCCACCGTTTTGGCCCTCCACAGATCTTCTTCAGAGTCCGTGC	384
		:::	:::
Dd	173	rAlaArgAlaAlaIalThnProThnPhe-----ProdlYCyStrhArgSerThnTrh	191
QY	383	ATCAAATCATTTATCAAGAGCATCTGGATGTGCAAAATTGATCTGCTTCCATCAAG--	329
		:::	:::
Dd	191	rAlaIenSer--SerArgSerArgASpArgSerThnCysArgProSerAlarArgGI	210
QY	328	----CCTTGGCCTTGTGCTGTCCGCTCAGACACTTGTAGAACCTGTCTGGCTCAGAGT	273
Dd	210	yPheProTrp-----CysArgProAlaPro-----	218
QY	272	ACACCTTCGCGTCCGCTGTTCACCTCGT-----	242
		:::	

```

Db      219 -ProArgAlaGlyAlaThrSerProProArgTrpProThrTrpCysSerArgArgPro 238
QY      241 -----TCGTCCAGCAACTCCAGTCTTCTCA 213
Db      238 gthraArgLeuProTrpAsnCysAlaAlaThrPheProAlaAlaProAsnAlaTrpVa 258
QY      212 CACACGCTGCACATCCCGCTGCGAGTCCCGCCCTTACAGGAAGGCGCGTGG 153
Db      258 lAlaArgArgAlaThrCysAlaCysArgAlaSerAlaCysThrTrpArgArgAlaAl 278
QY      152 TCGGACGCTGCGCTCCCGCGCGGCTCTCCGAGTCTTCCTCCCGCGAAGCCATC 93
Db      278 aArg-----ArgArgGlyLysSerSerCysArg-----ProTh 289
QY      92 TCGGCTCCGCTCCCGCGCATCGGCTCCCGCGCGC----- 56
Db      289 rArgAlaTrpThrAlaArgAlaSerSerProArgSerGlyArgCysSerAlaSerThrSe 309
QY      55 -----CCCTTCCCGCGGCGCGCGCGCGCGCGCGCAGC 24
Db      309 rAlaThrGlyArgTrpThrGlyArgPheProArgProThrCysArgArgArgTrpProth 329
QY      23 GCCGAGCTCGAGGCGACCCCA 2
Db      329 rProAlaAlaAlaProThrPro 336

RESULT 11
US-09-896-522-1 (1-1624) x US-09-896-522-1 (1-375)
/ Sequence 31128, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ PRIOR FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 31128
/ LENGTH: 375
/ TYPE: PRN
/ ORGANISM: Pseudomonas aeruginosa
US-09-896-522-1 (1-1624) x US-09-896-522-1 (1-375)

Alignment Scores:
Pred. No.: 2 41e-08 Length: 375
Score: 188.00 Matches: 112
Percent Similarity: 33.49% Conservative: 33
Best Local Similarity: 25.87% Mismatches: 133
Query Match: 6.34% Indels: 155
DB: 4 Gaps: 19

US-09-896-522-1 (1-1624) x US-09-896-522-1 (1-375)
QY      9 GCCTCCGACCTCGCGCTGCGCG-----GCCGCG 38
Db      27 AlAlaAspProGlyAlaGlyArgArgGlySerProGlyAlaCysProGlyAspArgArg 46
QY      39 CGGCGCGGGAAGGGCGCGCGCGGACCCGATCGCGCGGAGCGGA----- 86
Db      47 GluLeuGluGlyCysAlaLeuAlaSerGluProGlyProGlyAlaGlyCysAlaAlaVal 66
QY      87 -----GCCCGAGATGCTCGCGCGGAGGAACGACGCGAGCGCC 128
Db      67 ArgSerArgLeuArgThrGlyArgArgGlyAlaGlyGlyArgArgArgGlnArgPro 86
QY      129 CGCGCGCGGA-----GCCCGA-----CCGTCGCGACCGAGCGCCCTT 164
Db      87 AlAlaGlyPheProArgGlyGlnPheProArgProTrpProArgSerLeuSerGlyLeu 106

```

```

QY      165 CCGATAGGGGTGAGCGCGGCACTGCGAGCGGGAAGTGCACCGCTGTGTGAGAAGATCAT 224
Db      107 TrpArgArgGly----- 110
QY      225 GAGTGTCTGGGACAGAACAGGTGTGAGACGCGGACGGAAAGTGTATCTTGAGCCA 284
Db      111 -----AlaAspAlaThrArgGlyProGlyGlyLeuProArgGlyThrHisProGlyGly 128
QY      285 GGCACAGTTCTA-----CAAGTCTGACGGCAGA 314
Db      129 GlyGlnArgLeuLeuHisArgProAlaGlyGlyLeuPheArgProGlyArgProGlyArg 148
QY      315 GCAGAGGCGCAAGGCTTGAAGAGACATCAATTGACATCCAGATGCCTTGATTA 374
Db      149 ValGluGlyArgAla----- 153
QY      375 TGATTGTATGCACAGGACTCTGAGAACATCTGTGAGGCGAAGAAC-----GGTGAAGTGC 431
Db      154 -----AspArgAlaAspHisArgGlyGlyAlaHisValArgAlaGlyLeu 168
QY      432 GACCTTATGATTTGTGTACACACTCAAGGTTACAGAGACACAGGTGTCTACCTTCGGA 491
Db      169 Asp-----ProValGly 172
QY      492 CGTGTCTGTGTGAGGCACTGTGTGTCTACAGCCAGAGATCCGGGACATGTTCCA 551
Db      173 GlnGlySer-----ArgArgAlaGlyAsp----- 180
QY      552 CCGCGCCTCTGTTGAGACAGCACTCCGACGTCAAGCTGTCTGAGAAGTCTCCGGA 611
Db      181 -----ArgHeHisGlnValArgArgGlyAlaArgValProProArgAlaGly 196
QY      612 CGTGGCGCG-----AGGAGGACCTGTGAGCAAGATTCTGACGACATAC 656
Db      197 ArgArgGlnGlyGlyGlyAspArgArgAlaGlyProGlyProGlySerArgAlaLeuHis 216
QY      657 -----CACCTTCGTGAAGCGCGGCTTCGAGAGATTCGCGCTGC 695
Db      217 ArgGlyGlnGlyArgGlnCysHisProArgArgGlyGly-----ArgProAla 232
QY      696 GACAAAGATATGCGCATGT---GATCATCCACAGAGAGTGAACAATATGTTCCAT 752
Db      233 AspAspProAlaArgArgCysLeuArgHisProArgGlnAlaGlyAlaValTrpLeuGln 252
QY      753 CAACCTGATGTGCAGACACATCCAGACATTTGAATGTGTACATCTGCATAATGSCACG 812
Db      253 ArgArgGlnArgValGlyValProGly-----Leu 262
QY      813 AGGAGGTCCTCAATGGCGGAGC-----TACAAAGC 842
Db      263 ArgArgValGlnGlnAla-ProAlaValLeuProProLeuProAspGlyPheHisArgSe 282
QY      843 GACCTTTTTCGAGCCAGGGGACCACTCGGATGTGACCTCTGCAAAACGGTCAATT 902
Db      282 rSerAlaAspGlyProGlyThrGlnArgArgValGlyGlnGlnGlyProArgAlaHisArg 302
QY      903 GAGTCACAGACAGACCCCACTGAGGCGCTCGGACGCTCAGAGGAGGTCTCCCGCG 962
Db      302 gAlaThrAspProArg-----SerProAlaGln-AlaGlyGlyArgArgProG 317
QY      963 GCATGTGTGTTCAGGACTGAGCCTTGAGGACGCCCAACCAACCACTGCTCTCTCGG 1022
Db      317 LylLeuArgValArgProGlyGlyArgGlyAlaProLeuHisGlyLeu-----LeuSerGln 336
QY      1023 CGCAGCCCAAGGAGGTGTAGACAGAGGCTTCTCCTCACTCAGAGTGAAGAACTCAGATG 1082
Db      336 lArgProGlyGlyAspProArgArgArgArgLeuArgThrLeuArgSerGlnArgAlaArgP 356
QY      1083 TGTCACTCAGACTCAACTGTGAGCACTGAC 1115
Db      356 roSerSerArg-AlaPheSerTrpGlyThrGln 366

```



```

; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27679

Alignment Scores:
Pred. No.:      3,286-08      Length:      300
Score:          186.00      Matches:      107
Percent Similarity: 32.19%      Conservative: 15
Best Local Similarity: 28.23%      Mismatches: 106
Query Match:      6.31%      Indels:      153
DB:              Gaps:      18

US-09-896-522-1 (1-1624) x US-09-252-991A-27679 (1-300)

QY      1024 CGCCGAGAGAGACAGTGGGTGTGGGTGCGCTCCAGGCTCAATGCTCTGAACACACAT 965
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      7   ArgArgArgAlaGlyValGlyAlaGlyValArgValHisAlaArg----- 21

QY      964 GCGGGGGGAGACCTGCTGAGAGCTCGGAGAGCCCTCAGTGGGCTGTGCTGGACT 905
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      21 ----- 21

QY      904 CCAATGTGACGTTTGCCAGAGTGACATCCAGGGTGTGCTCCGTGCTCAAAAAG 845
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      22 -----ProAlaAlaArgSerArg 27

QY      844 TCCGCTTTAGTCCGCCCATTTGACCTCTCGGTGCCATTGGACAGATGCACATTCA 785
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      28   LeuAlaArgProAlaArgAlaGlyTyrThr-----GlyCybArgValAla 41

QY      784 GAATGTCCTGATGTGTGACGATGACATGATTGACACCATATTGTCCACTCTGTG 725
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      42   GlySerProAlaArgAlaAla-----ValProAlaAlaPro----- 53

QY      724 GGATGATCAGATCGGCATCTTCTTGTGCGAGCGAGCAATCTCTCGAAGCGGCTTCA 665
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      54 -----GlyArgArgProValArgAlaArg 62

QY      664 -----CGAAGGTGTGTACTGCGTCAAGATCTGCCAGGCTCTCC 623
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      63   GlyPheArgGlyProLeuArgArgSerCysAlaVal-----LeuProGlyThrSer 79

QY      622 CTGGGCGCAGCTGCCGAGAAC-----TCTTCGAGACAGCC 587
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      80   ProAlaThrValProProSerAlaAspGlyAlaTyrAlaProProSerAlaArgSerPr 99

QY      586 TGAAGTCGAGTGGTGTCCACGA-----AGAGGCGCAGGTGGAACATGTCC 540
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      99   O-----ValGlyCysHisProAlaGlyCysTyrSerAlaProGlySerThrThrAs 116

QY      539 CGGATCTCTGGCTGTGAACACCAAGA-----TG 510
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      116  pGlySerAlaGlyGlyArgSerProArgArgGlnGlyAlaAlaValThrProValProCy 136

QY      509 CCTCAAAACAGAACCCAGTCGCGAGGTAGACACCGTGTCTGTGTAACCTTGATGT 450
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      136  sProArg-----ArgThrValGlyGluPro-----Cy 145

QY      449 GTCAAAATCATAGTGGCAGCT-----CCACCGTTTGCCCTCC 408
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      145  sSerAlaGlyGlySerSerValProGlySerCysTyrArgSerValProGlyCysProGly 165

QY      407 ACGATGTTCTTCAGAGTCTGTGATCAATCATTAACAAAGCATGTGATGTCATAA 348
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      165  yArgCyseGlyGlnArgProAlaAlaProGly----- 175

QY      347 TTGACTGTCTTTCAAGGCTTGCGCTTGTGCTGCGGTCAGAGACTTGT-----AG 294
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      176  -----ProGlySerProCysCybArg 183

QY      293 AACCTGTCTGGCTCAGATGACACACCTTCGCTGCGCTGCTTCACCTGCTCTGTCCC 234
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      183  gThrCysProAlaProGlyArgArgSerAlaAlaAlaAlaThrAlaGlyArgThr---Gly 202

```

```

QY      233 AGCACTCCATGATCTTCTCACACACAGTGCATTCCTCCGTGGAGTGCCTGCTCAC 174
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      202  yAlaThrGlyTyrProSerGlyProGlyArgAspArgTyrArg-----Pr 218

QY      173 CCTATCAGAGAGGCGCGTGTGTC--GACAGTGTGCTCCGCGCGCGGCTCTCGAGTC 115
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      218  oAlaSerGlyAlaProArgGlyThrGlyArgSerAlaAlaGlyProGlyArgAlaAlaTr 238

QY      114 TTCGCTCCCGCGGAAGCATCTCGGCTCCGCTCCCGGCGCATGGGTCCCGCGCCGC 55
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      238  pAnArgPro-----SerSerAlaProProAlaGlyArg--ArgArgGlyAla 253

QY      54  CCGTCCCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      254  ProArgProGlyAlaGlyCysProAlaGlyArgCysGlyArgSerAlaPro 270

RESULT 14
US-09-252-991A-22452
; Sequence 22452, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: MARC J. RUBENFIELD et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ. ID NOS: 33142
; SEQ. ID NO 22452
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22452

Alignment Scores:
Pred. No.:      3,766-08      Length:      258
Score:          185.00      Matches:      92
Percent Similarity: 41.25%      Conservative: 14
Best Local Similarity: 35.80%      Mismatches: 83
Query Match:      6.28%      Indels:      69
DB:              Gaps:      18

US-09-896-522-1 (1-1624) x US-09-252-991A-22452 (1-258)

QY      691 GGCAGAACTCTCGAAGCGCGCTTACGAGAGTGTACTGCTCAGAAATCTGCTCCA 632
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      3   GlyArgGlyArgAspArgProGlyCysArg-----ProAlaPro 15

QY      631 GGTCCCTCCCTCGGCGCA-----CGTCCGAGAACTCTTCGAGACACCTGACGTGAGAT 575
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      16  ArgProAlaArgArgAlaArgArgProGlyArgArg-----ArgArgTyr 30

QY      574 CGGTGTCAAGAGAGCGCAGGTGGAACATGTCGCGATCTCTGCTGT---AGAAAC 518
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      31  ArgCysAlaThrAlaHisAlaAlaAlaArgArgAlaAlaAlaProGlyCysSerArgThr 50

QY      517 CCAGATGCCCTCAACACAGAACACAGTCCGAGGTGAGACACCGGTGTCTGTGTA--- 461
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      51  ProArgArgArgArgThrThrAlaArgAlaGlyArgAlaProAlaAlaArgAlaGly 70

QY      460 -----ACCTTGAAGTGTGCAAAATCATAGTGGCAGCT----- 425
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      71  SerProAlaGlyArgThrSerSerArgAlaValAlaGlyArgSerAlaProValArgCys 90

QY      424 -----CAACCGTTTGCCTCCACGA-----TGTTCT----- 398
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      91  SerAlaProProAlaValProAlaArgAlaAlaGlyArgAlaArgProCysSerLeuPro 110

```

QY 397 -----TCAGAGCTCTGCATCAATATATCAAGGACCTGGATGTCATAAATTGT 344
Db 111 ArgProGlyArgArgCysAlaSerArgArgGlyArgArg----- 123
QY 343 ACTGCTCTTCAAGGCTTGG-----CCTTCTGCTCCGCTCAGACCTTTGTAAGAC 290
Db 124 -----AlaAlaArgProGlyCysAlaProAlaAlaGlyLeuSer---ProArgThrPro 140
QY 289 TGTCTGCTCAGAGTGAACACCTTCGCTGCGC-----GCTGTTCGA---CCTGTTCT 239
Db 141 AlaProArgArg-----ProProAlaArgAlaArgGlnAlaProProAlaProArgArg 158
QY 238 GTCCAGAGACTCCATATCTTCTCAGACAGGTGACTTCCGCTGCACTGCCAGTCCGCC 179
Db 159 ArgProAlaValProGlySer-----TTPArgCysArgAla 170
QY 178 TCACCCCTATCAGAGAGCGCTGTGTGCGAGCGGTGCGCTCCGCGCGCGGCTGTGCG 119
Db 171 GlyProVal-----ArgAlaAlaProAla-----ProGlyArgGlyTyrArg 184
QY 118 AG-TCTTCGCTCCCGCGAAGCATCTCGGCTCCGCTCCGCGCATCGGATCCCGCG 60
Db 185 ArgArgSerProProAlaAlaAlaThrAlaProAlaAlaGlyAlaGlyArgThrSerAla 204
QY 59 ---CCCGCCCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGTGGGA 12
Db 205 GlyAlaAlaProAlaGlyArgProAlaAlaProProAlaAlaAlaProGly 221

RESULT 15

US-09-252-991A-31760
; Sequence 31760, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31760
; LENGTH: 1706
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31760

Alignment Scores:

Pred. No.: 9.63e-08 Length: 1706
Score: 184.50 Matches: 129
Percent Similarity: 32.82% Conservative: 41
Best Local Similarity: 24.90% Mismatches: 185
Query Match: 6.26% Indels: 163
DB: 4 Gaps: 25

US-09-896-522-1 (1-1624) x US-09-252-991A-31760 (1-1706)

QY 1388 AAAGTACTTGTTCACAAAACCTTTGAGTATGTGTGAGGTGTAAGAACAGATC--- 1332
Db 1230 LysAlaValArg11LevalThrArgGlnGlnAla1LealThrCysGluArgAspLeuGln 1249
QY 1331 AGATGGAATAAACTCTCTCCACTGTGGGTTCACGTCAACAATAATCAGGCGCAGCA 1272
Db 1250 ArgLeuAspArgAlaLeuAlaArgTyr-----GluAsnGlnAlaSerArg 1264
QY 1271 GTGTCTAGGCTGTCTCTCAATTTCCCAATATATGTCCTCACAAGAACCT 1212
Db 1265 LeuAlaGlnLeuSerAsp-----AlaGlnArgAlaAlaAlaHLeAla 1278
QY 1211 CCAGAGCTTCTGTGACATTTCTGTGTCATTTCTCAGTACCTAGAGGATCTTTAAACCGC 1152

Db 1279 ArgArgAlaSerLeuHisAlaLeuAlaArgThrGlyThrLeuAlaGlyArgAlaThrThr 1298
QY 1151 AACGAGCCTTAAGTGGCTGAAAACCTCAGAAAGCCTGTGCTAGTCCACAGATGTGATC 1092
Db 1299 GlyGlnAspArgGlyArg-11eProGlnAlaAspProAlaArgThrThrArgArgAlaArg-- 1317
QY 1091 TGAAGTACAC-----ATCTGAGTTTCCACTCTTGTAGTGAAGAGCCTCG 1047
Db 1318 -----HisProProGlyArgGlyLeThrProAlaThrProProProGlyLysAr 1335
QY 1046 CTGCTAACACTCCCTTGGGGTGGCGGAGAGAACGATGGGTGTGGTGGGCTCCCA 987
Db 1335 GglnArgProAlaProGlyProArgArgGlnAlaArgCysArgGlnArgGlySerProAl 1355
QY 986 GGCTCAGTCCCTGAACACACATCCGCGGCGGAGACCTGCTGAGGCTGCG----- 935
Db 1355 aaenProAla-----SerAlaGlyArgArgArgProProArgArgArgGlySerPr 1372
QY 934 -----CAGCCCTCAGTGGGGTGTGCTGTGAC 906
Db 1372 oAlaArgProArgPheCysArgProArgGlnArgGlnArgProArgArgGly-----Th 1388
QY 905 TCCAAATGTACCGCTTTGCCAGAGTCAAGCATCCAGGGTGTCCCTGCTCAGAAA 846
Db 1388 rPro-----GlnArgSerProAlaArgThrArgProAlaProGlnAspAr 1403
QY 845 GTCCGCTTGTAGTCCGCCCATTTGACACCTCTCGGCGCATTTGACAGATGTGACATTC 786
Db 1403 gArg-----AsnAlaHisHis-- 1408
QY 785 AGAATGCTGTGATGTGTGTCAGATCAGTATGATGATGATGATGATGATGATGATGAT 726
Db 1409 -----ProGlyThrValAlaAlaArgProPro----- 1416
QY 725 GGGATATCATCGGCATCTCTTGTGCGGACAGACATCTCTGA----- 677
Db 1417 -----AlaThrGlyArgProAlaArgThrAlaProG1 1427
QY 676 -----AGCGCGCTTCAAGAGTGGTGTACTGCGTCAAGTATGCTCCAGGCTCCTC 624
Db 1427 yAlaHisArgProAlaHisArgArgThrAlaThrAlaAlaArgArg----- 1442
QY 623 CTTGGCGCACGTCCTCCGAGAACTTTGACAGACAGCTGAGCGTGGATGCTGTCACAG 564
Db 1443 ---GlyGlnArgProSerValProArgAlaThrGlyThrArgArgSerArg----- 1458
QY 563 AAGAGCGCAGGTGGAACATGTCCCGGATCTCTGCTGTAGAACCAAGATGCCCTCA 504
Db 1459 -----ThrAlaProGlyAlaPro-----GlnProAlaAlaArgG1 1470
QY 503 AACGAGACACGTCGCGCAG---GGTAGACACCGGTGTCTGTGTAACCTTGAAGTGTTC 447
Db 1470 nProGlyProArgProArgProGlyArgPro----- 1480
QY 446 ACAAAATCATAGTGTGCGACCTTCACCGTTTGGCTCCACGATGTTCTTCAGAGTCTG 387
Db 1481 -ArgThrProAlaAlaThrProAlaProGlyThrProAlaArgProArgG1ArgSerAr 1500
QY 386 TGCATCAATCAT-----TATCAAGGATCTGATGTCGA 351
Db 1500 gArgProArgArgArgArgThrArgArgThrAlaProAlaArgArgArgLeuProAlaAr 1520
QY 350 AAATGTACTGTCCTTCAAGGCTTGGGCTTCTGCTGCGCTGAGACCTTGTGAAC 291
Db 1520 gGlnArgSerAlaAlaThrArgArgThrHisArgAlaMetGlnArgHisProHisArgPr 1540
QY 290 CTGTCTGTGCTCAGATGACACCTTCGCTGCGCTGCTGTTCACCTGTCTGTGTCCAGC 231
Db 1540 oProProAlaThrGlyArgPro-----ArgProProProGlyArgAlaProG1 1556
QY 230 AACTCATATCTTTCACACACAGGTGACCTTCGCGCTGAGAGTGCAGCGCTCACCCCT 171

```
Db      1556 Y---ProArgGlnProArgLeuArgSerProArgArg----- 1567
QY      170 ATCAGGAAGGCCCGCTGTGTCGGAC----- 146
Db      1568 -HISGlyHisArgLeuGlyAlaAspArgGlnGlyArgProAlaGlnThrArgHisProAr 1587
QY      145 -----GGTCGGCCTCCGGCGCGGGGC--TCTCG 120
Db      1587 gLeuArgProGlySerArgArgGlnGlyArgGlnArgProProAlaThrAlaArgArgAr 1607
QY      119 CAGCTTCGCGCTCCCGCGAAGCCATCTCGGCTCCGCTCCCGCGCATCGGGTCCCGCG 60
Db      1607 gProGlnArgGlnProArgGlnProAlaArgPro-ArgHisArg-----AsnProLeuV 1625
QY      59 CCGGCCCTTCCTCCCGGAGCCCGCGCGCCGCCCGCCAGCCCGAGTCGGAGG 10
Db      1625 aArgArgValProThrProProGlyProAlaArgArgProGlyArgArg 1641
```

Search completed: November 25, 2003, 07:42:32
Job time : 70.812 secs

This Page Blank (uspto)


```

XX WO20015312-A1.
PN
PD 26-JUL-2001.
PF 26-DEC-2000; 2000MO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
F1 Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
P1 Zhao QA, Zhou F, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
DR N-PSDB; AAI60444.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
PS the encoded polypeptides (AA038642-AA042213) with neurotropic,
XX immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
CC
CC
XX
XX Sequence 296 AA;
SQ
Alignment Scores:
Pred. No.: 2,26e-118 Length: 296
Score: 1551.00 Matches: 296
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 52.28% Indels: 0
Gaps: 0
US-03-896-522-1 (1-1624) x AAA41288 (1-226)
QY 37 GCCGCGCCCGGGGAAAGCGCGCGCGCGGAGCCCGATGCGCGGAGCGGAGCGGACCGGATG 96
Db 1 AAGAGTCTGATAGAGGCGTGAAGAGTCTGATAGAGTCTGATAGAGTCTGATAGAGTCTGATAG 20
QY 97 GCTTCGCGCGGAGCGGAGGAGTCTGATAGAGTCTGATAGAGTCTGATAGAGTCTGATAGAGTCTGATAG 15
Db 21 AAGAGTCTGATAGAGGCGTGAAGAGTCTGATAGAGTCTGATAGAGTCTGATAGAGTCTGATAGAGTCTGATAG 40
QY 157 CGGCGCTTCTGATAGAGGCGTGAAGAGTCTGATAGAGTCTGATAGAGTCTGATAGAGTCTGATAGAGTCTGATAG 21
Db 41 ArgProPheLeuIleGlyValSerGlyGlyTThrAlaSerGlyysSerThrValCysGlu 60
QY 217 AAGATCATGAGTCTGCGGAGCAGAGAGGAGTGAAGAGCGGAGCGGAGGAGTCTGATC 27

```

Ds	6	LysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValValIle
Qy	277	CTGAGCCAGAGACAGGTTCTACAAAGGTCCTGACCGCAGACAGCAAGAGCCAGGCTTGAAA
Ds	81	LeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeuLys
Qy	337	GGCAGATCAATTTTGGACCATCCAGATCGCTTTGTATGATTTGATGGCAGAGACTCTG
Ds	101	GLYGLNTrpAspPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeu
Qy	397	AAGACATCGTGGAGGCGCAAAAAGGTGGAGGTCGCCAGCTATGATTTTGTGACACTCA
Ds	121	LysAsnIleValGluGlyLysThrValGluValProThrTrpAspPheValThrHisSer
Qy	457	AGGTTACCAAGAACCAACGAGTGATCTACCCCTCGGACGCGGTCTCTGTTGAGGCACTTTG
Ds	141	ArgLeuProGluThrThrValValLysProAlaAspValValLeuPheGluGlyIleLeu
Qy	517	GTGTTCTACAGCAGAGAGATCCGGGACATGTTCCACCTGGCGCCTCTTCGTGACACCGAC
Ds	161	ValPheTrpSerGlnGlnIleArgAspMetPheHisLysLeuArgLeuPheValAspThrAsp
Qy	577	TCCGACGTGAGGCTGTCTGAAAGAGTCTCCGGACGTGCGCCGAGGAGAGGACCTTGAG
Ds	181	SerAspValArgLysSerArgTrpValLeuArgAspValAlaArgGlyArgAspLeuGlu
Qy	637	CAGATTCTGACGCAATCACCAACCTTCGTGAAAGCCGGGCTTCGAGAGAGTTCTGCTCCG
Ds	201	GlnIleLeuThrGlnIleThrThrPheValLysProAlaPheGluGlnPheCysLeuPro
Qy	697	ACAAAGAGATGAGCCGATGTCATCCACAGAGAGAGGAGCAATATGATGGTCCATCAAC
Ds	221	ThrLysLysTrpAlaAspValIleIleProArgGlyValAspAsnMetValAlaIleAsn
Qy	757	CTGATGTCGAGCAGACATCCAGGACATTCGATGGATGTCATCTGCAATGGCAGCCGAGA
Ds	241	LeuIleValGlnHisIleIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArgGly
Qy	817	GGGTCCAAATGGCGGAGCTTACAAAGGAGACCTTTTCTGAGCCAGGAGGACACCCCTGGATG
Ds	261	GlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGlyMet
Qy	877	CTGACCTCTGGGCAAAAGGTCACATTTGGAGTCCAGACAGACCCAC 924
Ds	281	LeuThrSerGlyLysArgSerHisLeuGlnSerSerArgProHis 296
RESULT 2		
AAAM39502	ID	AAAM39502 standard; Protein; 277 AA.
XX	AAAM39502;	
XX	AC	
XX	DT	22-OCT-2001 (first entry)
XX	DE	Human polypeptide SEQ ID NO 2647.
XX	XX	
KW	Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer	
KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;	
KW	leukaemia.	
XX	XX	
OS	Homo sapiens.	
XX	XX	
PN	MO20015312-A1.	
XX	XX	
PD	26-JUN-2001.	
XX	XX	
PF	26-DEC-2000; 2000MO-US34263.	
XX	XX	
PR	21-JAN-2000; 2000US-0488725.	
PR	25-APR-2000; 2000US-0552317.	

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AA892446 to AA895893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

SQ Sequence 277 AA;

Alignment Scores:

Pred. No.:	4.1e-110	Length:	277
Score:	1450.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	48.87%	Indels:	0
DB:	22	Gaps:	0

US-09-896-522-1 (1-1624) X AAB93941 (1-277)

QY	94	TTGGCTTGGGGGAGGAGCGAAGACTGCGAGAGCCCGGCGCGGAGGCGCAACGCTCCGAC	153
Dp	1	MetAlaSerAlaGlyGlyGlyAspCysGlnSerProAlaProGlnAlaAspArgTrpHis	20
QY	154	CAGCGGCGCTTCTGTATAGGGGTGAGCGGCGCGCACTCCAGCGGGAAGTCGACCGTGT	213
Dp	21	GlnArgProPheLeuIleGlyValSerGlyThrAlaSerClyLysSerThrValCys	40
QY	214	GAGAAAGATCATGAGTTCCTGGGACGAACAGAGTGGAACAGCGGACCGGAAGTGTCT	273
Dp	41	GluIysIleMetGlnLeuLeuGlnGlnAsnGlnValGlnGlnArgGlnArgLysValAla	60
QY	274	ATCCGAGGCSAGAGACGTTCTTACAAAGTSCCTGACGCGACAGAGAAAGGCAAGGCGCTTG	333
Dp	61	IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGlnGlnLysAlaLysAlaLeu	80
QY	334	AAAGACAGTACAAATTTTGACATCCAGATCCGCTTGTATATGATTTGATGACACAGACT	393
Dp	81	LysGlyGlnTyrAsnPheAspHisProSerAlaPheAspAsnAspLeuMetHisArgThr	100
QY	394	CTGAAGAACATCGTGGAGGGCAAAAGCGTGGAGAGTCCGCACTATGATTTTGTGACACAC	453
Dp	101	LeuLysAsnIleValGlnGlyLysThrValGlnValProThrLysAspPheValThrHis	120
QY	454	TCAAAGTTTACAGAGACACACGCGTGTCTACCCCTGCGGACGATGTTCTGTGTGAGGCATC	513
Dp	121	SerArgLeuProGlnThrThrValValTyrProAlaAspValValLeuPheGlnGlyIle	140
QY	514	TTTGTTCTTACAGCCAGAGAGATCCGGAGACATTTCCACCTCGCGCTCTTTCGTGACACC	573
Dp	141	LeuValIlePheTyrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr	160
QY	574	GACTCCGACGTAGAGGTGCTCGAAGATTTCTCCGGAAGTGGCGGAGGGAAGGACCTG	633
Dp	161	AspSerAspValArgLeuSerArgValValLeuArgAspArgAlaArgArgGlyArgAspLeu	180
QY	634	GAGCAGATTCTGACGCGAGTACACCACTTCGTGAAACCGGCTTTCAGAGAGTTCCTG	693

	AAE16592	AAE16592 standard; Protein; 277 AA.
Db	181	GIUGNIIeLeuThrdmIytrThrThrheValIysProAlaPheGIUGIunPheCysLeu 200
Qy	694	CGACAAAGAGTAGTATGCCATGTATCATCCACGAGAGTGGACATATATGGTTGCCATC 753
Db	201	ProThIyLysrIytrAlaIasPValIIeIlePProIrgIyAlaIasPaaMeIValAlaIle 220
Qy	754	AACCTATGTGCAGACATCCAGGACATTCGTGAATGGTGCACATGTGCMAATGGCACCGA 813
Db	221	AsnIeuIIeValGIhIhIerIleGIunIasPrlIeIuengIyAspIleCysIySTpPhIeArg 240
Qy	814	GAGGGGTCCAAATGGGGGAGGCTAACAAAGGGGACCTTTTGTGACGAGGGGACAAACCTGGG 873
Db	241	GIyGIySerAsnGIyArGIserGIyIyAsrIghIhrPheSerGIunProGIyAspHisPProGIy 260
Qy	874	ATTCCTAACCTGTGCAGAAACGATCACATTGTGGAGTTCAGACAGACACCCAC 924
Db	261	MetIeuThrSerGIyIyAsrGIserHisIleuGIunIserSerIAsrGIProHis 277

RESULT 4

ID AAE16592 standard; Protein; 277 AA.

AC AAE16592

DT 18-APR-2002 (first entry)

Human 57658 protein.

Human; uridine kinase-like protein; haematopoietic neoplastic disorder;

KW rheumatoid arthritis; psoriasis; dermatitis; osteoporosis; rickets;
KX protein; leukaemia; haemoglobinuria; diabetes mellitus
KY 57058

KW Alzheimer's disease: Parkinson's disease: Huntington's disease: KX bacillaria; myocardi; hyperextension; arteriosclerosis;

cytostatic: cardioprotective: neuroprotective: nootropic: gene therapy; chromosome mapping; tissue typing; dermatological; kw

anc1convulsant.

OS Homo sapiens.
 YX

FH	Key	Location/Qualifiers
ET	Mod:4:4:0d-0:4:0	E 10

FT	Residue	/label= N-myristoylation_site
25	151	

FT /note= "Kinase uridine transferase ATP-binding

cycle phosphopentokinase domain"

FT /note= "phosphoribulokinase domain; Protein kinase

FT	Modified-site	27..32
----	---------------	--------

FT	Binding-site	30..37
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

FT	Modified-site	34.36
----	---------------	-------

FT	Modified-site	
F1	38..41	/label= fluorocytosine

FT	Modified-site	100.0102
F1	/label=caseln_klnase_11_phosphorylacton_site	

ET	Modified-site	108-115	/label= Protein_kinase-c_phosphorylation_site
----	---------------	---------	---

FT	Domain	/label= tyrosine_kinase_phosphorylation_site
FT	Domain	154 277

/note= "Kinase uridine monophosphokinase transferase
ATP-binding kinase-like ribonucleoside pyrimidine ETS

FT	CDNA domain"
FT	100 100

```
FT      /label= Casein_kinase_II_phosphorylation_site
```

FT	/label= Protein_kinase-C_phosphorylation_site
1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	1
10	1
11	1
12	1
13	1
14	1
15	1
16	1
17	1
18	1
19	1
20	1
21	1
22	1
23	1
24	1
25	1
26	1
27	1
28	1
29	1
30	1
31	1
32	1
33	1
34	1
35	1
36	1
37	1
38	1
39	1
40	1
41	1
42	1
43	1
44	1
45	1
46	1
47	1
48	1
49	1
50	1
51	1
52	1
53	1
54	1
55	1
56	1
57	1
58	1
59	1
60	1
61	1
62	1
63	1
64	1
65	1
66	1
67	1
68	1
69	1
70	1
71	1
72	1
73	1
74	1
75	1
76	1
77	1
78	1
79	1
80	1
81	1
82	1
83	1
84	1
85	1
86	1
87	1
88	1
89	1
90	1
91	1
92	1
93	1
94	1
95	1
96	1
97	1
98	1
99	1
100	1

FT	/label= Protein_kinase-C_phosphorylation_site
1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	1
10	1
11	1
12	1
13	1
14	1
15	1
16	1
17	1
18	1
19	1
20	1
21	1
22	1
23	1
24	1
25	1
26	1
27	1
28	1
29	1
30	1
31	1
32	1
33	1
34	1
35	1
36	1
37	1
38	1
39	1
40	1
41	1
42	1
43	1
44	1
45	1
46	1
47	1
48	1
49	1
50	1
51	1
52	1
53	1
54	1
55	1
56	1
57	1
58	1
59	1
60	1
61	1
62	1
63	1
64	1
65	1
66	1
67	1
68	1
69	1
70	1
71	1
72	1
73	1
74	1
75	1
76	1
77	1
78	1
79	1
80	1
81	1
82	1
83	1
84	1
85	1
86	1
87	1
88	1
89	1
90	1
91	1
92	1
93	1
94	1
95	1
96	1
97	1
98	1
99	1
100	1

```
FT /label= N-myristoylation_site
```

FT	Modified-site	247..249	/label= Protein_kinase-C_phosphorylation_site
FT	Modified-site	251..254	/label= Casein_kinase_II_phosphorylation_site
FT	Modified-site	260..265	/label= N-myristoylation_site
FT	Modified-site	264..266	/label= Protein_kinase-C_phosphorylation_site
FT	Modified-site	264..267	/label= Amidation_site
FT	Modified-site	268..271	/label= Casein_kinase_II_phosphorylation_site
FT	Modified-site	273..275	/label= Protein_kinase-C_phosphorylation_site
XX			
PN	WO200202761-A2.		
XX			
PD	10-JAN-2002.		
XX			
PF	28-JUN-2001; 2001WO-US21063.		
XX			
PR	30-JUN-2000; 2000US-216503P.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
XX			
P1	Glucksmann MA;		
XX			
DR	WPI; 2002-140091/18.		
DR	N-PSDB; AAD27186.		
XX			
XX			
XX			
XX			
PS	Claim 4; Fig 1a; 103pp; English.		
XX			
CC	The patent discloses human uridine kinase-like polypeptides, designated		
CC	57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins		
CC	are useful for developing diagnostic and therapeutic agents for 57658-		
CC	mediated or related disorders such as haematopoietic neoplastic disorders		
CC	(e.g. leukaemia), haematopoietic disorders (e.g. psoriasis, dermatitis,		
CC	diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,		
CC	osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels		
CC	(e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's		
CC	disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or		
CC	their antibodies are useful in screening assays, detection assays (e.g.		
CC	forensic biology) and predictive medicine (e.g. diagnostic assays,		
CC	prognostic assays, and monitoring clinical trials and pharmacogenomics).		
CC	They are useful as reagents for diagnosing and treating 57658-mediated		
CC	disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,		
CC	to identify an individual from a minute biological sample (tissue typing),		
CC	and to aid in forensic identification of the biological sample. The		
CC	present sequence is human 57658 protein.		
XX			
SQ	Sequence 277 AA:		
Alignment Scores:			
Pred. No.:	4,1e-110	Length:	277
Score:	1450.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	48.87%	Indels:	0
DB:	23	Gaps:	0
US-03-896-522-1 (1-1624) x AAE16592 (1-277)			
QY	ATGGCTTCGGCGGAGCGCAGACTTCGAGAGCCCGCGCGCGAGCCGACCTCCGCAC	153	
DB	1 MetalaserIaAGlYgIyGluAaPcYsgIuSerProAlaProGluAaPaRgProHis	20	
QY	154 CAGCGCCCTTCTCTGATAGGGGTGAGCGCGGCGGACGACGCCAGCGGAAAGTCAGCTGTGT	213	
DB	21 GlnAGpProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys	40	

OY	214	GAGAAATCATGAGTGTCTGGGACAAAGAACGAGGTGAAACACGGGACACGGGAAAGTGTGTC	273
Db	41	GLULysIILeMeCluLeuLeuGLyGlnsngIuValGLuGlnAArgGlnAArgLysValVal	60
OY	274	ATCTGTAGCGAGGACAGGTTCTTACAAGGTCCTGACCGGACAGGACGAAAGGACCTTG	333
Db	61	ILeueSerGIInbPArgPheTyrlLysValLeuThrIlaGLuGlnLysValLysValLeu	80
OY	334	AAAGACAGTACATTTTGAACCATCCAGATGCGCTTTGATAATGATTTTGATGACAGACT	333
Db	81	LYSGIYLGIInTyAsnbnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr	100
OY	394	CTGAAGAACATCGTGGAGGGGCAAAACGGTGGAGGTCGCCGACCTTGTGAATTTTGGACACAC	453
Db	101	LeuLysAsnILeValGLuGLuLyLysThrValGLuValProIInTyAspPheValThrHis	120
OY	454	TCAAGGTACACAGACACACGGTGGTCTACCTCGGACGTCGTCTTGTGTTGAAGGCATC	513
Db	121	SeArArgLeuProGLuInThrThrValValTyrrProIlaAspValValLeuPheGLuGLyIle	140
OY	514	TTGGTGTCTTACACGCCAGAGATCCGGGACATGTTCCACTTGCGCCTCTTCTGTGGACACC	573
Db	141	LeuValPheTySerGIInLysIleArgAspMetPheHisLeuArgLeuPheValAspThr	160
OY	574	GACTCCGAGGTCAAGGCTGTCTGGAAGAGTCTCCGGGACGTCGGCCGAGGGGAGGACCTG	633
Db	161	AspSerAspAlaArgLeuSerArgArgValLeuAspAspValArgArgGLyArgAspLeu	180
OY	634	GACACAGATTCGTGACGAGTACACACACTTCGTGAAGCCGACCTTGACGAGATTCTGCTG	693
Db	181	GLuGlnILeLeuThrGLuInTyrrThrPheValLysProIlaPheGLuGLuPheCysLeu	200
OY	694	CCGACAAAGAGATATCCGATGTGATCATCCACGAGAGTGGACATATGTGTTGCCATC	753
Db	201	ProThrLysLysArgLysValAspValIleIleProArgGLyValAspAsnMetValAlaIle	220
OY	754	AACTGTATGCTGACGACATCCAGGACATTCGMAATGGTGCATCTGCAATGGCACACCGA	813
Db	221	AsnLeuILeValGLuHisIleGLuAspIleLeuAsnGLyAspIleCysLysThrPheArg	240
OY	814	GGAGGGTCCAAATGGGCGGAGTTCACAAGCGGACCTTTTCTGAGCGAGGAGACCACTGGG	873
Db	241	GLyGLySerAsnGLyArgSerTyrlLysArgThrPheSerGIInProGLyAspHisProGLy	260
OY	874	ATGCTGACTCTTGGCAACAGGTCAATTTGGAGTCCACACACAGACCCCCAC	924
Db	261	MetLeuThrSerGLyLysArgSerHisLeuGLuSerSerSerArgProHis	277
RESULT 5			
AAG64506			
ID	AAG64506 standard; Protein; 276 AA.		
XX	AAG64506;		
XX	AC		
DT	02-OCT-2001 (first entry)		
XX			
DE	Human uridine kinase.		
XX			
KM	Human; uridine kinase; UK.		
XX			
OS	Homo sapiens.		
XX			
PM	CNI287172-A.		
XX			
PD	14-MAR-2001.		
XX			
XX	07-SEP-1999; 99CN-0118818.		
XX			
XX	07-SEP-1999; 99CN-0118818.		
XX			
PA	(UTFU-) UNIV FUDAN.		

[illegible][illegible]

Db 61 ILeuSerGlnAspCysPheTyrIlyValLeuThrAlaGlnGlnValAlaLysAlaLeu 80
QY 334 AAAGACAGTAACTTTTGGACATCCAGATGCTTTGTAATGATTGATCCAGAGACT 393
Db 81 LysGlyGlnTyrAnPheAspHisProAspAlaPheAspAsnAspLeuMetHisLysThr 100
QY 394 CTGAAGAACATCGTGGAGGCAAAACGGTGGAGGTCGCGACTATGATTGTGACACAC 453
Db 101 LeuLysAsnIleValGlnGlyLysThrValGlnValProThrTyrAspPheValThrHis 120
QY 454 TCAAGTTACACAGACACACGAGTGTACCCCTGGAGCGAGTGTCTGTTGAGGGCATC 513
Db 121 SerArgLeuProGlnThrThrValValTyrProAlaAspValValLeuPheGlnGlyIle 140
QY 514 TTGGTGTTCACAGCCAGAGATCCGGACATGTTCCACTGCGCTCTTGTGAGACAC 573
Db 141 LeuValPheTyrThrGlnGlnIleAlaArgAspMetPheHisLeuAlaArgLeuPheValAspThr 160
QY 574 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGACGTCGCCCGAGGGAGGACCTG 633
Db 161 AspSerAspValAlaArgLeuSerArgArgValLeuAlaArgAspValGlnArgGlyArgAspLeu 180
QY 634 GAGAGATTCTGACGACATACACACCTTGTGAAGCGGCTTGAAGAGTTGCTGCTG 693
Db 181 GlnGlnIleLeuThrGlnTyrThrAlaPheValLysProAlaPheGlnGlnPheCysLeu 200
QY 694 CCGACAAAGAGTATGCCGATGTATCATCCACAGAGAGTGAACAATATGTTGCCATC 753
Db 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAlaAspMetValAlaIle 220
QY 754 AACCTATCTGCGACACATCCAGACATTTGTAATGTCATCTGCAATGGCACCGA 813
Db 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspLeuCysLysArgHisArg 240
QY 814 GGAAGGTCCATGGCGGAGCTTACAGAGCACTTTCTGACCGAGGAGCACCTGGG 873
Db 241 GlyGlyProAsnGlyArgAsnHisLysArgThrPheProGlnProGlyAspHisProGly 260
QY 874 ATGCTGACCTCTGGCAAAACGGTCACTTTGAGAGTCCAGACAGACCCAC 924
Db 261 ValLeuAlaThrGlyLysArgSerHisLeuGlnLysSerSerArgProHis 277
RESULT 8
ABB89353 ID ABB89353 standard; Protein; 190 AA.
AC ABB89353;
DT 24-MAY-2002 (first entry)
XX Human polypeptide SEQ ID NO 1729.
DE Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KM antileukemic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KM vulnereary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KM cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein.
OS Homo sapiens.
PN MO200190304-A2.
PD 29-NOV-2001.
PF 18-MAY-2001; 2001WO-US16450.
PR 19-MAY-2000; 2000US-205515P.
XX (HUMA-) HUMAN GENOME SCI INC.
PI Birse CE, Rosen CA;
XX WPI, 2002-122018/16.

DR N-PSDB; ABL89762.
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders -
XX Claim 11; SEQ ID NO 1729; 2081pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABL9449-ABL90853) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO.int/pub/publicated_pct_sequences.
XX
SQ Sequence 190 AA;
Alignment Scores:
Pred. No.: 2,56e-69 Length: 190
Score: 951.00 Matches: 187
Percent Similarity: 97.91% Conservative: 0
Best Local Similarity: 97.91% Mismatches: 3
Query Match: 32.05% Indels: 1
DB: Gaps: 0
US-09-896-522-1 (1-1624) x ABB89353 (1-190)
QY 94 ATGGCTTCGGCGGAGGCGAAGCATCGAGGCGCGCGGAGGCGGACCGTCCGCAC 153
Db 1 MetAlaSerAlaGlyGlnAspCysGlnSerProAlaProGlnAlaAspArgProHis 20
QY 154 CAGCGGCGCTTCTGTAAGGGGTGAGCGGCGGCACTCCAGCGGAGTGCACCTGTGT 213
Db 21 GlnArgProPheLeuIleGlyAlaSerGlyThr***SerGlyLysSerThrValCys 40
QY 214 GAGAAGATCATGGAGTTGCTGGAGCAAGAACAGAGTGAACAGCGGACGGAAAGTGTG 273
Db 41 GlnLysIleMetGlnLeuLeuGlnGlnAsnGlnValGlnGlnArgGlnArgLysVal 60
QY 274 ATCTGAGCAGAGACAGGTCTTCAAGGTCCTGACGCGAGAGAGAAAGCGGCTTG 333
Db 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGlnGlnLysAlaLysAlaLeu 80
QY 334 AAAGACAGTAACTTTTGGACATCCAGATGCTTTGTAATGATTGATCCAGAGACT 393
Db 81 LysGlyGlnTyrAnPheAspHisProAspAlaPheAspAsnAspLeuMetHisLysThr 100
QY 394 CTGAAGAACATCGTGGAGGCAAAACGGTGGAGGTCGCGACTATGATTGTGACACAC 453
Db 101 LeuLysAsnIleValGlnGlyLysThrValGlnValProThrTyrAspPheValThrHis 120
QY 454 TCAAGTTACACAGACACACGAGTGTACCCCTGGAGCGAGTGTCTGTTGAGGGCATC 513
Db 121 SerArgLeuProGlnThrThrValValTyrProAlaAspValValLeuPheGlnGlyIle 140
QY 514 TTGGTGTTCACAGCCAGAGATCCGGACATGTTCCACTGCGCTCTTGTGAGACAC 573
Db 141 LeuValPheTyrThrGlnGlnIleAlaArgAspMetPheHisLeuAlaArgLeuPheValAspThr 160
QY 574 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGACGTCGCCCGAGGGAGGACCTG 633

DT 31-JUN-2001 (first entry)
XX Human transferase HTFS-1, SEQ ID NO:1.
XX
XX Human transferase; HTFS, agonist; antagonist; cellular signalling;
KM proliferation; cell proliferative disorder; immune disorder;
KM atherosclerosis; hepatitis; psoriasis; cancer; tumour;
KM inflammation; AIDS; Addison's disease; allergy; asthma; anaemia;
KM cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus;
KM multiple sclerosis; rheumatoid arthritis; pancreatitis;
KM systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;
KM haemodialysis; extracorporeal circulation; trauma; transgenic animal;
KM gene therapy; drug screening.
XX Homo sapiens.
XX WO200132888-A2.
XX 10-MAY-2001.
XX 02-NOV-2000; 2000WO-US30485.
XX 04-NOV-1999; 99US-0163595.
XX (INCY-) INCYTE GENOMICS INC.
XX Tang YT, Yue H, Hillman JL, Lal P, Bandman O, Patterson C;
PI Shih IL, Azamzai Y, Lu DAM, Baughn MR; Patterson C;
XX WPI; 2001-328796/34.
XX N-PSDB; AAH23801.
XX Human transferase polypeptides and polynucleotides useful for
PT diagnosis, prevention and treatment of cell proliferative and immune
PT system disorders and for identifying agonists and antagonists -
XX Claim 1; Page 103-104; 157pp; English.
XX Sequences AAB73494-AAB73535 represent novel human transferase proteins
XX HTFS-1 to HTFS-42, and sequences AAH23801-AAH23842 represent cDNAs
XX encoding them. The proteins play important roles in the regulation of
XX cellular signalling and proliferation. The HTFS proteins are useful for
XX screening compounds for their effectiveness as agonists or antagonists of
XX transferase activity, or for compounds that specifically bind to an HTFS
XX protein or which modulates the activity of an HTFS protein.
XX Pharmaceutical compositions comprising an HTFS protein, HTFS
XX agonist or antagonist, or genetic construct encoding an HTFS
XX protein are useful for treating a disease or condition associated
XX with decreased or increased expression of functional HTFS. Disorders
XX which may be treated using such compositions include cell proliferative
XX disorders and immune disorders. For example, diseases which may be
XX treated include atherosclerosis, hepatitis, psoriasis, cancers (including
XX breast, bladder, bone marrow, brain and uterus cancer), inflammation,
XX AIDS, Addison's disease, allergies, asthma, anaemia, cirrhosis, Crohn's
XX disease, atopic dermatitis, diabetes mellitus, multiple sclerosis,
XX rheumatoid arthritis, pancreatitis, systemic lupus erythematosus,
XX thrombocytopenia, and ulcerative colitis. They may also be used to treat
XX complications of cancer, haemodialysis, extracorporeal circulation,
XX trauma and haemopoietic cancer, including lymphoma, leukemia and
XX myeloma. Polynucleotides encoding HTFS proteins are useful for creating
XX transgenic animals to model human diseases, for diagnostic purposes and
XX to generate hybridisation probes useful in mapping the naturally
XX occurring genomic sequences. HTFS, and its catalytic or immunogenic
XX fragments are useful for screening libraries of compounds in a variety of
XX drug screening techniques. Antibodies which specifically bind HTFS may be
XX used for the diagnosis of disorders associated with the expression of
XX HTFS, or in assays to monitor patients being treated with HTFS or
XX agonists, antagonists or inhibitors of HTFS. The present sequence
XX represents an HTFS protein of the invention.
XX Sequence 261 AA;

Alignment Scores:

	Pred. No.:	7.05e-68	Length:	261
Score:	934.00	Matches:	189	
Percent Similarity:	79.41%	Conservative:	27	
Best Local Similarity:	69.49%	Mismatches:	40	
Query Match:	31.48%	Indels:	16	
DB:	22	Gaps:	4	
US-09-896-522-1 (1-1624) x AAB73494 (1-261)				
QY 115	GACTGCGAGAGCCCCGGCGGAGGCGGACCGTCCGAC--CAGCGGCGCTTCTCGATA	171		
Db 4	AspSerGlnGlnThrLeuGlnAsnHsGlnGlnProAsnGlyGlnGluProPheLeuIle	23		
QY 172	GGGATGAGCGGCGGACCTGCGACGCGGAGTGCACCGTGTGTGAGAAATCATCATGAGTT	231		
Db 24	GlyValSerGlyGlyThrAlaSerGlyLysSerValCysAlaIleValIleLeuSerGlnAspSer	43		
QY 232	CTGGGACAGAACGAGGTGAAACAGCGGACGGAAGGTGTCATCTTGACCGACAGCAGG	291		
Db 44	LeuGlyGlnAsnGlnValAspTyrArgGlnLysGlnValIleValIleLeuSerGlnAspSer	63		
QY 292	TTCTACAGGTCTCTGACGCGACGACGAGGAGCCAGGCGCTTGAAAGACATGATTTT	351		
Db 64	PheTyrArgValLeuThrSerGlnGlnLysAlaLysAlaLeuLysGlyGlnPheAsnPhe	83		
QY 352	GACCATCCAGATGCGCTTGTGATTAATGATTTGATGACACGACCTGTAAGAAACGTGGAG	411		
Db 84	AspHisProAspAlaPheAspAsnGlnLeuIleLeuLysThrLeuLysGlnIleTrGlu	103		
QY 412	GGCAAAACGGTGGAGGTCGCGACCTATGATTTGTGACACACTCAAGATTATCCAGACG	471		
Db 104	GlyLysThrValGlnIleProValTyrAspPheValSerHisSerArgLysGlnIleThr	123		
QY 472	ACGGTGTCTACCTTCGCGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	531		
Db 124	ValThrValTyrProAlaAspValAlaLeuPheGlnGlyIleLeuAlaPheTyrSerGln	143		
QY 532	GAGATCCGGGACATGTTCCACCTGCGCTTGTGTGACACGACCTCCGACGTCAGGCTG	591		
Db 144	GlyValAlaArgAspLeuPheGlnMetLysLeuPheValAspThrAspAlaAspThrArgLeu	163		
QY 592	TCTGGAAGATTCTCGGAGCGTG--CGCGAGGAGGAGGACCTGGAGGATTTCTGACG	648		
Db 164	SerArgArgValLeuArgAspIleSerGlnArgGlyArgAspLeuGlnIleLeuSer	183		
QY 649	CATTAACACACCTTCTGTGAAGCGGCTTGTGAGAGATTCTGCTGCCGACAAAGAGTAT	708		
Db 184	GlnTyrIleThrPheValLysProAlaPheGlnGluPheCysLeuProThrLysTyr	203		
QY 709	GCGATGTGATCTCCACGAGGAGTGACCAATATGTTGCCATCAACCTGATCGTGACG	768		
Db 204	AlaAspValIleIleProArgGlyAlaAspAsnLeuValAlaIleAsnLeuIleValGln	223		
QY 769	CACATCCAGACATCTTGAATGTGACATCTGCMAATGCGACCGAGAGGTCATCATGCG	828		
Db 224	HisIleGlnAspIleLeuAsnGly--GlnGly 232			
QY 829	CGAGGTACAGGCGACCTTTCTGTGACGAGGAGACCCCTGGATGCTGACCTTGCG	888		
Db 233	ProSerLysArgGlnThr--AsnGlyCysLeuAsnGlyTyrThrProSerArg 249			
QY 889	AAAGGTCACTTGTGATTCGACGACGACGACCCAC 924			
Db 250	LysArgGlnAlaSerGlnSerSerSerArgProHis 261			
RESULT 11				
ID ABP41393	standard; Protein; 337 AA.			
AC ABP41393;				
XX 22-AUG-2002 (first entry)				
DT				
XX				

DE Human ovarian antigen HOELP29, SEQ ID NO:2525.
 XX
 KM Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KM ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KM infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KM PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KM inflammatory condition; immune disorder; blood disorder;
 KM cardiovascular disorder; respiratory disorder; neurological disorder;
 KM gastrointestinal disorder; urinary system disorder; drug screening;
 KM gene therapy; chromosome mapping; forensic analysis;
 KM antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KM antiinflammatory; gynaecological; reproductive; chromosome 1p32.
 OS Homo sapiens.
 XX
 XX MO200200677-A1.
 XX
 PD 03-JAN-2002.
 XX
 PF 07-JUN-2001; 2001MO-US18569.
 XX
 PR 07-JUN-2000; 2000US-209467P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 XX WPI, 2002-147878/19.
 XX N-PDB; ABO54470.
 DR
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT isolated in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX
 PS Claim 11; SEQ ID NO 2525; 29222p; English.
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABO54131-ABO56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosting or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX
 SQ Sequence 337 AA;

Alignment Scores:
 Pred. No.: 1.64e-67
 Score: 930.00
 Percent Similarity: 73.86%

Length: 337
 Matches: 201
 Conservative: 25

Best Local Similarity: 65.69% Mismatches: 57
 Query Match: 31.34% Indels: 23
 DB: 23 Gaps: 6
 US-09-896-522-1 (1-1624) x ABP41393 (1-337)
 QY 12 TCAGACCTGGCGCTTGGCGGCGCGCCCGGGAGAGGGCGCGGACCGG 71
 Db 54 SerApsAlGIYAlaGly--GlyGluArgAlaSerValArgThrGlySerGly----- 70
 QY 72 ATGCGCGGAGCGGAGCGCGAGATGGCTTCGCGGAGCGGAGAGACTGGAGACCCCGC 131
 Db 71 ---ArgArgGlyGlyAlaAsnHisGlyArgGlyGlnArgAlaAspProAlaGluProPro 89
 QY 132 GCCGAGGCGCGACCGTCCGACACGAGCGCCCTTCCTGATAGGGGTAGCGGCGGCTGC 191
 Db 90 AlaAlaGlnArg-----ArgArgAlaLeuProTyArgArgHis-GlyGlyThrAl 106
 QY 192 CAGCGGAGATCGACCGTGTGAGAAAGATCATGAGTGTGCGGACAGACGAGGTGA 251
 Db 106 AserGlyLysSerSerValCysAlaLysIleValGlnLeuLeuGlyGlnAsnGluValAs 126
 QY 252 ACAGCGGACGCGAAGAGTGTATCTCTGAGCGAGACAGATTCTACAGGCTCTGACGCGC 311
 Db 126 pTyArgGlnLysGlnValIleLeuSerGlnAspSerPheTyArgValLeuThrSe 146
 QY 312 AGAGCAGAAAGCGCCAGGCTTGAAAGACAGATTCATTTGACATCCAGATGCTTTGA 371
 Db 146 rGluGlnLysAlaLysAlaLeuLys**GlnPheAsnPheAspHisProAspAlaPheAs 166
 QY 372 TATATGTTGATGACACAGACTCTGAAAGAACTCGTGAGCGCAAAACGATGAGGTGCC 431
 Db 166 PsnGlnLeuIleLeuLysThrLeuLysGluIleThrGlnGlyLysThrValGlnIleLeu 186
 QY 432 GACCTATGATTTTGTGACACACTCAAGTTTACAGAGACACAGGTGTACCTGCGGA 491
 Db 186 oValTyArgPheValSerHisSerArgLysGlnGluThrValThrValTyProAlaAs 206
 QY 492 CGTGTTCTGTTTGAAGGACATCTGGTGTCTTACAGCCAGAGATCCGGACATGTTCCA 551
 Db 206 pAlaValLeuPheGlnGlyIleLeuAlaPheTySerGlnGlnValArgAspLeuPheGln 226
 QY 552 CCGGCGCTCTCGTGGACACCGACTCGACGCGGCTGTCTCGAAGATCTCCGGGA 611
 Db 226 nMetLysLeuPheValAspThrAspAlaAspThrAspLeuValLeuArgAs 246
 QY 612 CGTG--CCCGAGGAGGAGACCTGAGGACAGATTGTGACGACATCACCACTTCGTGA 668
 Db 246 pIleSerGluArgGlyArgAspLeuGlnIleLeuSerGlnTyIleThrPheValLys 266
 QY 669 GCCGCGCTTGAAGAGTTTGTGCTGCGCCGACAAAGATATGCCGATGTATATCCACG 728
 Db 266 sProAlaPheGlnGlnPheCysLeuPheProThrLysTyAlaAspValIleLeuProAr 286
 QY 729 AGGAGTGCACATPAGTGTGACATCAACTGATTCGCGAGCAGCATTCAGACATTTGAA 788
 Db 286 gGlyAlaAspAsnLeuValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAs 306
 QY 789 TGGTGACATCTGCMAATGACACCGAGAGGTTCATATGGGCGAGCTACAGCGGACCTT 848
 Db 306 nGly-----GlyProSerLysArgGlnThr-- 314
 QY 849 TTCTGAGCCAGGGGACACCTGGAGTGTGACCTTGCGAAACGATTCATTTGGAGTC 908
 Db 315 -----AsnGlyCysLeuAsnGlyTyThrProSerArgLysArgGlnAlaSerGlnSe 932
 QY 909 CAGCAGCAGACCCAC 924
 Db 332 rSerSerArgProHis 337
 RESULT 12
 AAB56582
 ID AAB56582 standard; Protein; 337 AA.

XX	AAB56582;
AC	
XX	
DT	13-MAR-2001 (first entry)
XX	
DE	Human prostate cancer antigen protein sequence SEQ ID NO:1160.
XX	
KW	Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KM	neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;
KV	vulnerable; gastrointestinal; nephrotoxic; antineoplastic; gynaecological,
KW	antibacterial; gene therapy; neutral; immune; reproductive; renal;
KX	gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX	wound; infectious disease.
OS	Homo sapiens.
XX	
PN	WO20005174-A1.
PD	
XX	21-SEP-2000.
PF	
PR	08-MAR-2000; 2000MO-US05988.
XX	
XX	12-MAR-1999; 99US-0124270.
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	(ROSE/) ROSEN C A.
PI	Rosen CA, Ruben SM;
XX	
DR	WPI; 2000-587513/55.
N-PSDB; AAF15785.	
PT	Prostate cancer associated gene sequences, referred to as prostate
PT	cancer antigens, useful for treatment, prevention, and diagnosis of
PT	disorders such as prostate cancer -
PS	Claim 11; Page 1566-1567; 2338pp; English.
XX	
CC	AAF15566 to AAF16505 encode the human prostate cancer associated
CC	proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC	The prostate cancer antigens can have neuroprotective, cytostatic,
CC	cardioactive, immunomodulatory, muscular, vulnerable, gastrointestinal,
CC	nephrotoxic, anti-infective, gynecological and antibacterial activities,
CC	and can be used in gene therapy. The prostate cancer antigen
CC	polynucleotides may be used for detection of prostate cancer, chromosome
CC	identification, as chromosome markers, and for numerous other diagnostic
CC	or research purposes. The prostate cancer antigens may be used to treat
CC	disorders such as neural, immune, muscular, reproductive,
CC	gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC	disorders, wounds, and infectious diseases. AAF15506 to AAF16514 to
CC	AAB57303 represent sequences used in the exemplification of the present
XX	invention.
XX	
SQ	Sequence 337 AA;
XX	
Alignment Scores:	
Pred. No.: 4,21e-67	Length: 337
Score: 925.00	Matches: 200
Percent Similarity: 73.53%	Conservative: 25
Best Local Similarity: 65.36%	Mismatches: 58
Query Match: 31.18%	Indels: 23
DB: 21	Gaps: 6
US-09-896-522-1 (1-1624) x AAB56582 (1-337)	
OY	12 TCGAGCTTCGCGCCTTGGCGGCGGCCCGCGGGAAGGCGCGCGCGGAGCCCG 71
Dd	
54 SerAPAlagIyAlagIy---GIyGuArGaAlaseRValArghrGlySerGIy----- 70	
OY	72 ATGCCCGGAGCGGAGCGCGAATGGCTTCGCGGCGGAGCGGAAGACTGCAGAGCCCCCG 111
Dd	
Db	71 --AtGARgGIyGIyAlaasnhtsgIyArGgIyGlnArGAlAsProlAgInuProfo 89

QY	132	GGCGGAGGCGGACCGTCCGACCCAGCGGCCCTTCTGATAGGGGTGACGGCGGCACTGC	191
Db	90	AlaAlaGlnArg-----ArgArgAlaLeuProTyrArgArgHis-GlyGlyThrAl	106
QY	192	CAGCGGGAGAGTCGACCGGTGTGTGAAGAATCATGTGAGTCTCTGGACAGAACGAGTGA	251
Db	106	AsSerGlyHisSerSerValCysAlaTyrIleValGlnLeuLeuGlyGlnAsnGluValAs	126
QY	252	ACAGGGCAGCGAGAGGTGTGTCATCTGAGCCAGAGCAGAGGTTCTTACAGAGTCTTCAGCGC	311
Db	126	PTyrArgGlnIysGlnIleValIleLeuSerGlnAspSerPheTyrArgValLeuThrSe	146
QY	312	AGACGACAGAGGCCAGGCTTGAAGAAGACAGTACAAATTTTGACCATTCGACATGCTTTGA	371
Db	146	rgIuGlnIysAlaIysAlaLeuIys***GlnPheAsnPheAspHisProAspAlaPheAs	166
QY	372	TAATGATTGTGACACGACTCTGGAAGAACATGCTGAGAGGGCAAAACGGTGGAGGTGC	431
Db	166	pasnGlu***IleLeuYsrThrLeuYsgIuIleThrGluGlyIysThrValGlnIlePr	186
QY	432	GACCTATGATTGTTGTGACACACTCAAGTTACAGAGACACGAGTGTCTTACCTCGGGA	491
Db	186	oValTyrAspPheValSerHisSerArgIysGluGluThrValThrValTyrProAlaAs	206
QY	492	CGTGTTCTGTTTGAAGGCATCTTGTTGTCTTACAGCCAGAGAGATCCGGGACATGTTCCA	551
Db	206	pValValLeuPheGlnIuGlyIleLeuAlaPheTyrSerGlnGluValArgAspLeuPheG	226
QY	552	CCTCGCGCTCTCTGTGACACCGCACTCCGACGTCAGCGTGTCTGGAAGAGTTCCTCGGGA	611
Db	226	nMeIlyLeuPheValAspThrAspAlaAspThrArgLeuSerArgArgValLeuArgAs	246
QY	612	CGTG---CGCGAGGAGGAGACCTGAGCAGATTTCTGACGCACTACACCACTTCGTGA	668
Db	246	pIleSerGlnArgGlyIyrArgAspLeuGlnIuIleLeuSerGlnTyrIleThrPheValIy	266
QY	669	GCCGGCCTTCGAGGAGTTCGTGCTGCGCCGACAAAGAGTATGCCAGTGTATCATCCACG	728
Db	266	sProAlaPheGlnIuGlnPheCysLeuProThrIlySlyTyrAlaAspValIleIleProAr	286
QY	729	AGAGGTGACAAATATGGTGTGCATCAACCTGATCGTGTGACACATCCAGGACAAATTCGAA	788
Db	286	gGlyAlaAspAsnLeuValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAs	306
QY	789	TGTGTGACATCTGCMAATATGGACACCGAGAGGGGTCCAAATGGCGGAGTCAACAGCGGACCTT	848
Db	306	ngIy-----GlyProSerIlyAspGlnIhr--	314
QY	849	TTCTGAGCCAGGAGCAACCTCTGGATGTCTGACCTCTGGCAAAACGTCACATTTGAATC	908
Db	315	-----AsnGlyCysLeuAsnGlyTyrThrProSerArgIlyArgGlnAlaSerGlnSe	332
QY	909	CAGCAGAGACCCCGAC 924	
Db	332	rSerSerArgProHis 337	
RESULT 13			
ID	AAW40480	standard; Protein: 335 AA.	
AC	AAW40480;		
XX	22-OCT-2001 (first entry)		
DE	Human polypeptide SEQ ID NO 5411.		
XX	Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoradiation;		
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;		
XX	leukaemia.		

OS	Homo sapiens.
XX	
PN	W0200153312-AL.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000MO-US34263.
XX	
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0598042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
P1	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
P1	Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
P1	Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX	
DR	WPI: 2001-442253/47.
DR	N-PDSB; AA159636.
PS	
XX	Example 2; SEQ ID NO 5411; 10078bp; English.
CC	The invention relates to human nucleic acids (AA157798-AA161369) and
CC	the encoded polypeptides (AA138642-AA42213) with nootropic,
CC	immunosuppressant and cytoskeletal activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localised neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilisation of the activities such as: Immune system suppression,
CC	CC Activin/inhibin activity, chemotactic/chemokinetic activity, hemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemias and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
XX	specification.
XX	
SQ	Sequence 335 AA;
Alignment Scores:	
Pred. No.:	3,34e-66 Length: 335
Score:	914.00 Matches: 203
Percent Similarity:	72.36% Conservative: 30
Best Local Similarity:	63.04% Mismatches: 57
Query Match:	30.81% Indels: 32
DB:	Gaps: 8
US-09-896-522-1 (1-1624) x AAA40480 (1-335)	
OY	8 CGCCTCGACCTGCGGGCTGGGGCGGGGCCCGGGAGAACGGGCGGGCGGA 67
Dd	31 ArgThrArgProSerAlaProA--AlaProSerValGlyCyGlyLylArArgArglu 49
OY	68 CCCGATCCGGGGAGCGGAGCGCCGAGATGGCTTCGGCG-----G 106
Dd	50 SerAspAla-GlyAlaGlyGly-GluArgAlaSerValaArgThryIyserylArArg 69
OY	107 GAGGCGGAAGACTCGAGAGCCCCGCGCCGAGACCGTGCAGCACGAGG----- 158
Dd	69 IyglArghTrhMetAlaGlyApbserGlunInThrleudnIasnhIsIndInProng 89
OY	160 -----CCCTTCGTGATAGGGGTTAGCGCGGCACTTGCCAGCGGGAAGTCGACCCTGT 211

Accession	Protein	Position	Sequence	Length
D8	1yGlyInpRhpheulleGlyValSerGIyThrlAserGIyLysSerValC	109		109
QY	212 GTGAMAGATCATGGAGTGTCTGGACAGAACGAGGTGACAGCGGACGGAGAGTGG	271		271
Db	109 ySalalalytlEValdInleuEnuGIyGlnAsnGIValAspIryrArgGlnIyGlnValY	129		129
QY	212 TCATCTGAGCCAGGACAGATTCTACAAAGTCTTGAACGCGACAGACGAAGCCAGGCT	311		311
Db	129 alIEleuSerGIInAspSerPheTyIryrGValIleuthrSerGIInIySalalalySal	149		149
QY	332 TGAAGGACAGAACATTTTGAACCTCCAGAGCGCTTGTATATGATTTGATCAGAGA	350		350
Db	149 euLySGIyGlnPheAsnPhesphIspCoAspAlaPheAsnphngIuIleIeLyuIst	169		169
QY	392 CTCTGAAMACATCGTGGAGGGCAAAACGCTGAGGTCGCGACCTTATGATTTTGTACAC	451		451
Db	169 hrIeulLyGIuIlehrGlnIyLysIthrValGIInIleProValIryrAspPheValSerh	189		189
QY	452 ACTCAAGGTTACAGAGACCAACGAGTGCTTACCTCGCGAGCTGTCTGTTTGAGGGCA	511		511
Db	189 IsSerArGIyGIuIInhrValIThrValTyProIalAspAlValIeupheGIuGIyI	209		209
QY	512 TCTTGAGTGTCTTACAGCCAGGAG--ATCCGGACATGTTCCACCTCGGCTTTGTGG	568		568
Db	209 IeLeuAlaPheTyIerGIInGIuArGIleArAspLeupheGIIneItyLysLeuPheValA	229		229
QY	569 ACACCGACCTCGACGTCAAGGCTGTCTCGAAGATTCTCCGGACGCG--CGCCAGGGA	625		625
Db	229 spIhrAspAlaAspIhrArGIeuserArGIyValIeulysAspIleSerGIuArGIyA	249		249
QY	626 GGGACCTGAGAGCATTTCTGACGACATCAC--ACCTTGTGAAGCCGGCTTGGAG	682		682
Db	249 rGIAspLeuGIuGIInIleuSerSerSerThrIeulArGheValIySpProIalPheGIuG	269		269
QY	683 AGTTTGCTGCTCCGCAAAAGATATGCCGATGTATATATCCACAGAGATGACAAATA	742		742
Db	269 IuPheCyLeuPheProIyLysIryrAlaAspValIleIeProArGIyAlaAspAsnA	289		289
QY	743 TGGTGCATCAACCTGATCGTGCGACATCCAGACATTCGTGAATGGTGACATCTGCA	802		802
Db	289 rGIValProIeAsnIeulIeValGIInhIsIleGIInpIleuAsnGIy-----	305		305
QY	803 AATGGACCGAGGAGGTCCATATGGCGGAGCTACAGCGGACCTTTTCTGACCGAGG	862		862
Db	306 -----GlyProSerAsnArGIInhr-----AsnGIyC	315		315
QY	863 ACCACCTGGAGATGCTGACCTCTGGCAAAACGCTCACTTTGAGATCCAGACGACCCC	922		922
Db	315 ySLeulAsnGIyTyThrProSerArGIyArGIInAlaserGIuSerSerArGIyProh	335		335
QY	923 AC 924			
Db	335 Is 335			
RESULT 14				
ABB62307				
ID	ABB62307	standard; Protein; 260 AA.		
XX	ABB62307;			
XX	AC			
XX	DT	26-MAR-2002 (first entry)		
XX	DE	Drosophila melanogaster polypeptide SEQ ID NO 13713.		
XX	KW	Drosophila; developmental biology; cell signalling; insecticide;		
XX	XX	pharmaceutical.		
XX	OS	Drosophila melanogaster.		
XX	PN	WO200171042-A2.		
XX	PD	27-SEP-2001.		

XX 23-MAR-2001; 2001WO-US09231.
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI, 2001-6656860/75.
 DR N-PSDB; ABL06410.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PT
 XX
 PS Disclosure; SEQ ID NO 13713; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 260 AA;

Alignment Scores:
 Pred. No.: 7,84e-52 Length: 260
 Score: 738.00 Matches: 143
 Percent Similarity: 82.94% Conservative: 32
 Best Local Similarity: 67.77% Mismatches: 34
 Query Match: 24.87% Indels: 2
 Gaps: 2

US-09-896-522-1 (1-1624) x ABB62307 (1-260)

QY 160 CCCTTCCTGATAGGGGTGAGCGCGGCGACTGCCAGCGGAAGTGCACCTGTGTGAGAG 219
 Db 27 ProtheneulleglyValAlaGlyGlyThrAlaSerGlySerThrValCysIleValys 46
 QY 220 ATCATGAGTGTCTGGACAGAAACGAGGTGGAACGCGGCGAGCGGAGGTGTCATCTG 279
 Db 47 IleMetGluGlnLeuGlyGlnAlaGluMetAspHisThrGlnArgGlnValIleSerIle 66
 QY 280 AGCCAGAGACAGGTCTCAAGGTCTTCAAGCGGACAGCAAGGCGCAAGCCTTGAAGA 339
 Db 67 SerGlnAspSerPheIleArgGluLeuThrProAlaGluIleValAlaGlnIleValysGly 86
 QY 340 CAGTACATATTTGACATCCAGATGCTTGTATATATTGATGACAGAGCTGCTGAG 399
 Db 87 LeupheusnheasphsIleProAspAlaPheAsnGluLeuMetCysSerThrLeuGln 106
 QY 400 AACATCTGAGAGGGCAAAACGGTGAAGTGCAGCTGATATTGTGACACACTCA-- 456
 Db 107 AsnIleLeuYegIleHisIleValGluIleProSerIleAspIleValGlnIleVal 126
 QY 457 AGGTACCAAGAGACAGCGGTGCTTACCTCCGAGCTGCTGCTTGAAGGCACTTGG 516
 Db 127 AspPheGluAsnValLeuValIleIleTyrProAlaAspValValLeuPheGluGlyIleLeu 146
 QY 517 GTGTTTACACCGACGAGATCCGGGACATGTTCCACTGCGCTTCTGTGACACCGAC 576
 Db 147 ValPheIleTyrPheProIleIleArgGluLeuPheHisIleTyrIleValIleAspThrAsp 166
 QY 577 TCCGACGTACGGCTGTCTCGAAGAGTCTCCGGGACGTG--CGCCAGAGGAGGACCTG 633

Db 167 SerAspThrArgLeuAlaArgValProArgAspIleAsnGluIleValArgIleAspLeu 186
 QY 634 GAGCAGATTCTGACGAGTACACCACTTGTGTGAAGCGGCTTCCAGAGCTTCTGCTG 693
 Db 187 AspAlaValLeuThrGlnIleValMetThrPheValIleProArgGlyAlaAspAsnThrValAlaIle 206
 QY 694 CCGACAAGAAGTATGCGATGATCATCCACGAGAGTGCACATATGTTGCATC 753
 Db 207 ProThrIleYsPheAlaAspValIleIleProArgGlyAlaAspAsnThrValAlaIle 226
 QY 754 AACCTGATCGTGACAGCATCCAGACATCTG 786
 Db 227 AspLeuIleValIleHisIleGlyGluIleLeu 237
 RESULT 15
 ABG27216
 ID ABG27216 standard; Protein; 326 AA.
 AC ABG27216;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #27207.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Dmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS91403.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PT
 PS Claim 20; SEQ ID No 57575; 103pp; English.
 XX
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 326 AA:

Alignment Scores:

Pred. No.:	2,33e-41	Length:	326
Score:	610.50	Matches:	143
Percent Similarity:	59.52%	Conservative:	32
Best Local Similarity:	48.64%	Mismatches:	68
Query Match:	20.58%	Indels:	51
DB:	22	Gaps:	9

US-09-896-522-1 (1-1624) x ABG27216 (1-326)

```
QY 7 TCGCCTCCGACCTCGGCGCTGGGCGG-----CGC 36
Db 26 SerProGlySerGluArgLeuAlaGlyAspGlyHisValArgGluProArgGluArgLeu 45
QY 37 GCCGCGCCCGGAGAGGCGCGCGCGCGCGGACCGCATGC--GCGGAGCGGAGCGCGAG 93
Db 46 AlaSerAlaAlaGluSerGlyGlySerProThrArgAlaArgAlaGlySerValArgPro 65
QY 94 ATGGCTTCGGCGGAGCGGAGGAGCTGCGAGAGCCCGCGCG--GAGCGCGACCGTCCG 150
Db 66 PheAlaGlnAlaAlaGlyGlyAlaAlaGluProTrpProGlyThrAlaSerArgPro 85
QY 151 CACGAG-----CGGCCCTTCCTGATAGGGGTGAGCGGCGGC 186
Db 86 CysArgThrThrSerSerProThrAlaAlaSerThrSerLeuIleGlyValThrTrpGly 105
QY 187 ACTGCCAGCGGAGAGTCGACCGTGTGTGAGAGATCATGAGTGTCTGGAGCAAGAG 246
Db 106 ThrAlaSerGlyLysSerSerValCysAlaAlaIleValGlnLeuLeuGlyGlnAnglu 125
QY 247 GTGGAACAGCGGAGCGGAGAGTGTCTATCTGAGCCAGAGAGTTCTACAAAGTCTG 306
Db 126 ValAspTyrArgGlnArgGlnValAlaIleLeuSerGlnAspSerPheTyrArgValLeu 145
QY 307 ACGGCGAGAGCAAGGCGGAGGCTTGAAGAGCAGTACATTTTGACCATTCAGATGCC 366
Db 146 ThrSerGlnGlnLysAlaLysAlaLeuLysGlyGlnPheAsnPheAspHisProAspAla 165
QY 367 TTGTATATGATTTGATGCACAGACTCTGAGAAACATCGTGAGGCGCAAAACGTGAG 426
Db 166 PheAspAsnGlnLeuIleLeuLysThrLeuLysGlnIleThrGlnGlyLysThrValGln 185
QY 427 GTGCCGACTATGATTTGTGACACTGACACTGATACAGAGACACGCGTGTACCTT 486
Db 186 IleProValTyrAspPheValSerHisSerArgLysGlnGluThrValThrValTyrPro 205
QY 487 GCGGAGCTGTGTTGTTGAGGGCATCTTGTTCTACAGCCAGAGATCCGGGACATG 546
Db 206 AlaAspAlaValLeuPheGlnGlyIleLeuAlaPheTyrSerGlnGlnValArgAspLeu 225
QY 547 TTCACCTGCGGCTCTCTGTGACACCGACTCCGACGTCAGGCTGTCTGAGAGA--GTT 603
Db 226 PheGlnMetLysLeuPheValAspThrAspAlaAspThrArgLeuSerArgArgGlyIle 245
QY 604 CTCGGGAGCTGCGCGGAGGAGGAGCTGAGCAGATT-----642
Db 246 MetAsnLeuLeuHisProArgGlyLeuArgAlaIleThrIleAlaValPheGlyLys 265
QY 643 -----CTGACGCACTACACCACTTCCTGTAAGCCGGCC-----675
Db 266 GlnAsnThrTyrIleArgLeuGlnProPheArgIleAsnValProProThrIleThrLys 285
QY 676 ---TTCGAGAGGTTCC--TGCCTGCCG-----696
Db 286 HisIleGlnGlnLeuGlnCysAlaProValValThrGlnLeuSerArgLysGlnGluThr 305
QY 697 ---ACAAAGAGATGCGGATGTGATCATCCACGAGAGATG 735
Db 306 ValThrValTyrProAlaAspValValLeuPheGlnGlyIle 319
```

Search completed: November 25, 2003, 07:30:34
Job time : 86.9182 secs

This Page Blank (uspto)